



Digitális tartalmak azonosítása, hitelesítése, hiteles tartalomszolgáltatás

SÜTHEŐ Péter

Loca Credibilia – az internet hiteleshelye

A Loca Credibilia elnevezés a magyar történelem ismerőinek a közjegyzőség középkori előképét jelenti. Hogy kerül egy XIII. századi intézmény neve egy internetes szolgáltatásra? A XI–XII. században, az írásbeliség elterjedése előtt egy-egy vitás ügy – például birtokper – esetén elkerülhetetlen volt, hogy a tanúságtevőket minden esetben személyesen is meghallgassák és e meghallgatások eredményeként szülessen ítélet. Az információ (tanúság) hitelessége tehát azon múlt, hogy annak a forrása (tanúságtevő) a vizsgálat időpontjában az adott helyen mit állított. Az egyszer hitelesnek elfogadott információ újra felhasználására, más időpontban hitelesnek történő elismerésére nem volt mód, mindig új tanúságtételre volt szükség. Az információ hitelességének attribútumai: a forrás, a tartalom és a közzététel időpontja térben és időben elválaszthatatlanok voltak az ellenőrzés helyétől és idejétől. Először az írásbeliség elterjedésével született meg az az intézmény, amelyik alkalmas volt arra, hogy az egyszer hitelesnek tartott tanúságtételt és annak idejét rögzítse, és pecsétjével igazolja, azaz lehetővé tette, hogy az ellenőrzés helyén és idejében ne kelljen megjelennie a tanúságtevőnek, helyette egy pecsétes írás igazolja, hogy a múltban tett tanúságát hitelesnek lehet elfogadni. Az ezt tartalmazó irathoz hozzárendeltek egy „adatot”, a hiteleshely

pecsétjét, ami a hitelességet igazolta. A XXI. századi hiteleshely is hasonlóan működik: hozzárendel a digitális dokumentumokhoz egy „pecsétet”, amely arra hivatott, hogy jelezze, a dokumentum hiteles. A bizalmi szolgáltatások (trusted services) fejlesztése a modern informatika egyik legfontosabb, leggyorsabban fejlődő területe. Nem minden bizalmi szolgáltatást hívunk azonban hiteleshelynek, sőt. A Loca Credibilia nevet csak a magyar feltalálók szabadalommal védett eljárása használja (EP3491575). Az elnevezés egyrészt tisztelgés a középkori előkép előtt, másrészt megkülönböztetés más bizalmi szolgáltatásoktól. A XXI. századi hiteleshely ugyanis – más szolgáltatásoktól eltérően – mindazon adatok tárhelye, amelyek segítségével egy digitális dokumentumpéldány hitelessége ellenőrizhető. A hitelesítés során tároljuk a forrás azonosító adatait, a dokumentum tartalmának digitális ujjlenyomatát és a hitelesítés időpontját. A dokumentumpéldányon pedig elhelyezünk egy pecsétet, ami olyan egyedi azonosítót tartalmaz, amely logikailag megbonthatatlan módon rendeli egymáshoz a dokumentumpéldányt és a tőle fizikailag elkülönítve tárolt hitelesítő adatokat. A „pecsét” tehát már többet jelent, mint a középkorban. Ott pusztán annak autoritása biztosította az irat hitelességét, ma pedig a pecsét az ellenőrizhetőségen, ilyen módon a megbízhatóságon alapuló hitelesség eszköze is.

Mi szüksége van egy könyvtárnak hiteleshely szolgáltatására, hiszen a könyvtár még a digitális rengetegben is per definitionem hiteles forrásnak minősül? Az információkereső olvasók éppen azért részesítik előnyben egy könyvtár digitális állományát egy tetszőleges weboldallal szemben, mert bíznak a könyvtári dokumentum hitelességében. Miért kell ide egy harmadik fél pecsétje? A rövid válasz az, hogy a könyvtári dokumentum hitelessége addig ellenőrizhető, amíg az állományban fellelhető – lásd középkori tanúságtétel az írásbeliség elterjedése előtt. Azaz, a letöltés pillanatában még biztosan hiteles, de később ezt a hitelességet már könnyű kétségbe vonni, hiszen egy digitális dokumentum szándékosan vagy sem, de minden eddigénél könnyebben módosítható. Érdekl ez a könyvtárat? Igen, ha a digitális információforrások versenyében meg akarja őrizni ma még kitüntetett helyét.

Az Országos Széchényi Könyvtár (OSZK) Országos Könyvtári Rendszer (OKR) projektjében az infrastrukturális fejlesztések mellett olyan kutatás-fejlesztési részprojektek is megvalósultak, amelyek a könyvtárak, közgyűjtemények szakmai feladatainak

magasabb szintű ellátását hivatottak támogatni. Ezek közé tartozik a Loca Credibilia rendszer prototípusának kifejlesztése is. A prototípusfejlesztés első szakaszában egy kutatásban vizsgáltuk, alkalmas-e egyáltalán a Loca Credibilia rendszer logikája a könyvtári felhasználásra. Az alábbiakban az ezt igazoló kutatásjelentésnek a legfontosabb megállapításait tesszük közzé.

A kutatás tárgya, módszertana

A Loca Credibilia hitelesítési eljárása két szakaszra bontható. Az első szakaszban azonosítani kell a hitelesítendő objektumokat, a másodikban pedig képezni azokat az adatokat, amelyek alapján a hitelesség ellenőrizhetővé válik. A prototípusfejlesztés során kutatást végeztünk annak megállapítására, hogy milyen elvek alapján célszerű a könyvtári dokumentumok, dokumentumpéldányok azonosítását elvégezni. A kutatás kezdetén a következő kérdéseket tettük fel:

1. Mi az azonosítás tárgya egy digitális gyűjteményben? Hogyan definiálható pontosan a digitális objektum fogalma?
2. Az azonosításnak milyen kritériumai érvényesülnek egy digitális gyűjteményben?
3. Milyen azonosító rendszereket használnak a digitális gyűjtemények?
4. Lefedik-e ezek az azonosítók a digitális gyűjteményekből származó dokumentumok hitelesség-ellenőrzésének minden igényét?

Hipotézisünk az volt, hogy a 4. kérdésre a válasz nemleges lesz, ezért azt is vizsgálat tárgyává tettük, milyen azonosító rendszer lenne alkalmas a feladat ellátására.

A kutatás első része a szakirodalom feldolgozására irányult. Áttekintettük a könyvtár- és információtudományi szakirodalmat, és a releváns publikációk alapján rögzítettük a szükséges fogalmak szakmai konszenzuson alapuló definícióit. A jelentés tartalmazza a feldolgozott szakirodalom bibliográfiai hivatkozásait. A második részben a „best practice” vizsgálat eszközével éltünk, megkerestük azokat az azonosító rendszereket, amelyek a legelterjedtebbek a digitális gyűjtemények körében és elvégeztük azok összehasonlító vizsgálatát. Az elemzés elsődleges célja az volt, hogy megállapítsuk: a prekonceptciónak megfelelően tényleg létezik-e olyan valós igény a digitális gyűjtemények kezelése kapcsán, amelynek ezek a rendszerek nem felelnek meg. A kutatás záró szakaszában pedig olyan logikai modelleket építettünk fel, amelyek alapján kialakítható egy olyan azo-

nosító rendszer, amely képes a nem lefedett igények kielégítésére is.

A digitális objektum fogalma

A digitális könyvtár fogalma valamikor az 1990-es évek közepén keletkezett. Szakmai (politikai?) legendák szerint az elnevezés *Al Gore*-hoz, az USA alelnökéhez köthető (Yin-Leng Theng, 2002). Való igaz, 1993 előtt a szakirodalomban hiába is keressük ezt a kifejezést, helyette viszont megtaláljuk az „elektronikus” és a „virtuális” jelzőket. Az „elektronikus könyvtár” szókapcsolat elterjedését leginkább a *The Electronic Library* című folyóirat 1983-as megjelenéséhez köthetjük, míg a „virtuális könyvtár” esetében nem találunk pontosan kijelölhető évszámot” (Sütheő, 2001, 211.p.). Elmondhatjuk azonban, hogy az azóta eltelt mintegy 20–25 évben sem sikerült egyértelműen és mindenki számára megnyugtatóan definiálni egyik fogalmat sem. A melléklet számos, a nyomtatott szakirodalomban és mérvadónak tekinthető internetes oldalakon megjelenő definíciót tartalmaz. Ezekből leginkább az szűrhető le, hogy a „digitális könyvtár” két diszciplína, a könyvtártudomány és az informatika metszéspontjába tartozik, ebből fakadóan legalább kettős a megközelítése, és több aspektusból le lehet és le kell írni. Ezek közül az egyik az intézményi, *institucionális* megközelítés, amely szerint a digitális könyvtár a hagyományos könyvtár funkcióit hivatott megvalósítani digitális gyűjteményeken (pl. Belkin, 1999.). A másik megközelítés *technikai*, amelyik arra helyezi a hangsúlyt, hogy a „digitális könyvtár” egy információs rendszer vagy információkereső rendszer, tehát alapvetően informatikai eszköz, amelynek célja, hogy a felhasználókat a számukra releváns tartalmakkal lássa el, teljesen függetlenül attól, hogy ezek a tartalmak hol helyezkednek el (pl. ARL, 1995.). Egy harmadik megközelítés pedig, a „digitális” szóra fókuszálva a tartalmak *tárolási és szolgáltatási formátumaira* és azok hordozóira helyezi a hangsúlyt (i2010).

A definíciók abban legalább megegyeznek, hogy a digitális könyvtár, hagyományos testvérétől eltérően, digitális formában rögzített dokumentumokat őriz. A „digitális dokumentum” fogalma azonban már nem ennyire egyértelmű. A definíciókat böngészve az olvasónak kicsit az az érzése, mintha a könyvtárosvilág nem tudna mit kezdeni a „digitális”, az informatikai világ pedig a „dokumentum” fogalmával. Való igaz, hogy a digitális dokumentumok egy sor olyan

sajátossággal rendelkeznek, amelyeket a hagyományos megközelítéssel nehéz értelmezni.

Első sajátosság a formátum. A digitális dokumentumok elektronikus formában léteznek. A legegyszerűbb meghatározásuk sokszor annyi, hogy nem papír alapon vagy nem nyomtatott médiumként. Ez a komplementerként történő meghatározás persze erősen pontatlan, hiszen a „nem nyomtatott” kategóriába ugyanúgy beleférnek a kéziratos művek (pl. kódexek) is, mint az elektronikusok, de a „nem papír alapú” megkötés is tág teret enged, érthető akár pergamenre, agyagtáblára vagy bármely más, az emberi szem számára látható jelek rögzítésére alkalmas hordozók használatára. Ezekről azonban könnyen belátható, hogy nem részei – legalábbis eredeti formájukban – egy digitális könyvtárnak. Ha nagyon pontosan akarunk fogalmazni, akkor az „elektronikus formátum” megfogalmazásba is bele lehet kötni. A digitális ellentéte ugyanis nem a „nem elektronikus”, hanem az analóg. Bekerülhetnek-e ennek alapján egy digitális könyvtárba analóg rögzítésű elektronikus dokumentumok? Jelen dokumentumnak azonban nem célja a terminológiai kérdések tisztázása, tehát megelégedhetünk annyival, hogy a digitális könyvtárban „digitális dokumentumokat” őriznek. Ezeket elektronikus formában rögzítik. A hordozó lehet offline (CD, DVD stb.) vagy online elérhető. Abban azonban egyetérthetünk, hogy a digitális könyvtár elsősorban az online elérhető tartalmakra fókuszál (lásd leginkább: ARL, 1995.).

A formátumot nemcsak a rögzítés módja, hanem a tartalom is meghatározza. Digitális dokumentumok tartalmazhatnak szöveges, képi, de hangzó és mozgóképi elemeket, sőt ezek kombinációit is. Ez innentől fájlformátumok sokaságát jelenti, megfejeelve azzal, hogy a számítástechnika világában meglehetősen régen jelen van az „osztott tartalom” fogalma. Ennek első megfogalmazásaként *Vannevar Bush* MEMEX vízióját (Bush, 1945), a gyakorlatba való átültetés első példájának *Ted Nelson* XANADU projektjét (Nelson, 1965.), széles körben való elterjedéseként pedig *Tim-Berners Lee* ötlete nyomán a World Wide Webet szokás emlegetni (Berners-Lee, 1992).

Az online elérhető, összetett és osztott tartalom alapvetően más megközelítést kíván a gyűjtemény kezelőjétől, mint a hagyományos („nem elektronikus”) könyvtári dokumentumok világában. Az első különbség mindjárt a dokumentum határaitra vonatkozik. Az világos, hogy már egy összetett dokumentum sem biztos, hogy megfeleltethető egyetlen fájlnak, osztott dokumentumok esetében pedig még az sem

biztos, hogy a fájlok egy helyen (pl. azonos szerve-
ren) található. Világos tehát, hogy a hagyományos
könyvtári dokumentumtipológia kevés támpontot
nyújt, ha korrekt és általános érvényű definíciót ke-
resünk a digitális dokumentumokra.

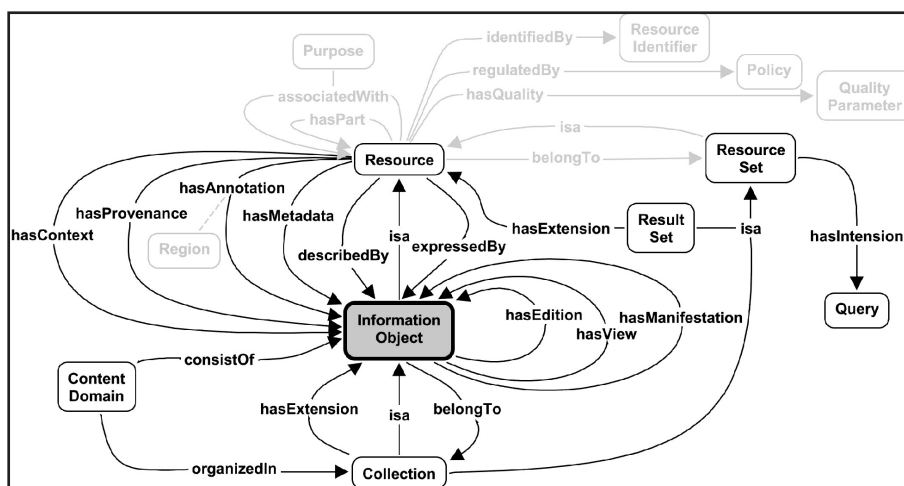
Ha a szakirodalomban a „digitális dokumentum”
definícióját keressük, nem is találunk még annyira
sem egységes megközelítést, mint a „digitális könyv-
tár” esetében. A bizonytalanságokat legtalálóbban
Buckland fogalmazza meg. Szerinte „a technológiára
helyezett hangsúly akadályozza a digitális dokumen-
tumoknak mint dokumentumoknak a megértését”
(*Buckland*, 1997).

A digitális könyvtár referenciamodelljei

Ahhoz, hogy a digitális dokumentum, digitális ob-
jektum fogalmát a könyvtártudományi terminológia
szerint is értelmezni tudjuk, célszerű megvizsgálni
azokat a modelleket, amelyek a gyűjtemények felépi-
tését ábrázolják. Számos ilyen modell létezik, de szá-
munkra két megközelítés bír kiemelt relevanciával.
Az egyik a DELOS munkacsoport által kifejlesztett
modellnek a „DL.org. Coordination Action on Digital
Library Interoperability, Best Practices & Modelling
Foundations” projektben, az Európai Unió 7. kutatási
keretprogram támogatásával továbbfejlesztett mo-
dellje, amelyre a továbbiakban DLRM-ként hivat-
kozunk, a másik pedig az IFLA 1998-as irányelveire
(A bibliográfiai tételek funkcionális követelményei –

Functional Requirements for Bibliographic Records
(FRBR)) és az arra épülő, 2016-ban draft verzióként
publikált FRBR Library Reference Model (a továb-
biakban FRBR LRM), a Könyvtári referenciamo-
dell. Utóbbi kifejezetten nem digitális gyűjtemények
modellje, hanem a hagyományos könyvtári gyűjte-
ményekre vonatkozó modell, de egyrészt tartalmaz
„távoli elérésű elektronikus gyűjtemények”-re vonat-
kozó megállapításokat is, másrészt mivel a digitális
könyvtár egy speciális könyvtárként is értelmezhető,
nyilvánvaló, hogy arra is alkalmazható.

Belkin például a digitális könyvtárat úgy definiálja,
mint olyan intézményt, ami a könyvtár funkcióit va-
lósítja meg digitális formában osztott és hálózatba
kapcsolt „információs objektumok” gyűjteményei-
nek kontextusában (*Belkin*, 1999.) Az „információs
objektum” kifejezést átvették a DLRM alkotói is.
Eszert tehát az „információs objektum” a „forrás”
(resource) fogalmába tartozik, annak kifejeződése.
„Az információs objektum ábrázolhat bármilyen in-
formáció egységet, legyen az szöveges dokumentum,
kép, hangzó dokumentum, multimediális dokumen-
tum vagy 3D objektum, beleértve a játékokat és a
virtuális valóság dokumentumait, valamint adatsoro-
kat és adatbázisokat. Az információs objektum ezen
felül tartalmazhat összetett objektumokat, valamint
információs objektumok gyűjteményét is.” (DLRM,
39.p.). Az információs objektum ezen felül még egy
sor tulajdonsággal rendelkezik, amelyeket az alábbi
ábra is tartalmaz.



1. ábra
A DLRM modell

Az „információs objektum” tehát meglehetősen elvont, de pontosan definiált kifejezés, amelyet azután az absztrakció különböző szintjein lehet értelmezni (DLRM, 40. p.). Az „információs objektum” egyrészt olyan „forrás”, amelynek van (vagy lehet) metaadata, kapcsolódhat hozzá annotáció, meghatározható az eredete, lehet kontextusa. De az „információs objektumnak” lehetnek kiadásai, nézetei, megjelenései, amelyek ugyanúgy információs objektumok. Az információs objektumnak számunkra különösen fontos jellemzője, hogy fölérendelt fogalmától, a forrástól örökli az azonosítottságot, mint tulajdonságot („identified By” ⇔ Resource Identifier; lásd az 1. ábrán!).

Az információs objektumokat tovább lehet specifikálni elsődleges céljuk szerint. Ennek értelmében beszélhetünk:

- *elsődleges információs objektumokról*, mint pl. önálló könyvek vagy adatsorok,
- *metaadat típusú információs objektumokról*, amelyek az elsődleges objektum leíró adatai,
- *annotáció típusú információs objektumokról*, amelyek a célforrás (általában elsődleges információs objektum) interpretálásában segítenek.

De specifikálhatók a kódolásuk szempontjából is, azaz lehetnek

- *ember által értelmezhetőek*, mint pl. szöveg vagy kép,
- *gép által értelmezhetőek*, mint pl. adatbázisok táblái vagy adatsorai, ERM modellek vagy ontológiák formális struktúrái.

Keletkezési körülményeiket és a valóságban fellelhető objektumokhoz való viszonyukat tekintve pedig lehetnek

- *eleve digitálisan keletkezettek*,
- *nem digitális eredeti digitalizációja útján keletkezettek*,
- *más információs objektumokról rögzített meta-adatok*, függetlenül attól, hogy az eredeti digitális vagy nem digitális formában áll-e rendelkezésre.

Az információs objektum fizikai megjelenése tehát egyértelműen és mindig digitális. A szerzők egyébként az „*információs objektum*” kifejezést a korábbi modellekben használt „*digitális objektum*” kifejezés helyett használják ebben a modellben (DLRM, 63. p.). Nem tévedünk tehát nagyot, ha azt mondjuk, hogy a köznapi használatban a „*digitális objektum*”, vagy „*digitális dokumentum*” fogalmakat a további-

akban ezen referenciamodell „*információs objektum fogalmával*” azonosítjuk, és annak tulajdonságaival dolgozunk.

Amikor tehát a kutatás a digitális objektumok azonosításának problémáját tűzi ki célul, akkor valójában a digitális könyvtár referenciamodellje szerint a források és azok bármilyen absztrakciós szinten megjelenő kifejeződéseinek, az információs objektumoknak az azonosításával kell foglalkoznunk.

Az FRBR modell, mint az információs objektumok értelmezési modellje

Az FRBR (Functional Requirements for Bibliographic Records), azaz a bibliográfiai tételek funkcionális követelményei az IFLA által kidolgozott ERM (Entity-Relationship-Model) modellje, amely a bibliográfiai tételekről gyűjtött, vagy a bibliográfiai tételekben fellelhető információkat ábrázolja olyan szempontból, hogy azok milyen módon érdekelhetik a felhasználót, a keresések során milyen módon tudják a felhasználói igényeket a legpontosabban kielégíteni. Az FRBR modell tehát nem a digitális könyvtárak számára kifejlesztett modell, alapvető célja nem is azok modellezése, vagy iránymutatás azok kialakításához, hanem minden könyvtári gyűjtemény számára elméleti modellként kíván szolgálni a céljainak megfelelő tárgyban. Egyértelmű az is, hogy megállapításai elsősorban a hagyományos könyvtári fogalmakon és kategóriákon nyugszanak. Maga a munka is a 90-es években zajlott, és a zárójelentés, ami a modellnek az IFLA grémiumai által elfogadott változatát tartalmazza, 1998-ban jelent meg (Functional Requirements for Bibliographic Records. Final Report. IFLA Series on Bibliographic Control 19. München: K.G. Saur Verlag, 1998.) Ez a munka tehát – formálisan mindenképpen – függetlenül és jórést korábban zajlott, mint a fent vázolt Digital Library Reference Model (DLRM) kidolgozása. Az FRBR modell tehát nem a DLRM modellből való levezetéssel készült, a DLRM sem az FRBR modell absztrakciója, mégis vannak közöttük olyan megfelelések, amelyek fontosak számunkra, ha a digitális objektumok azonosításának kérdésében tisztán akarunk látni. A rokonságot a DLRM egyébként több, az FRBR-ra tett hivatkozással, esetenként kritikával is jelzi (DLRM 40. p.).

Az FRBR modell 10 entitást különböztet meg és ezeket három csoportba sorolja. Az első csoportba tartozó entitások azok, amelyek a legnagyobb rokonságot mutatják a DLRM „információs objektum”

entitásával. Úgy is felfoghatjuk őket, mint az információs objektum különböző absztrakciós szinteken történő értelmezéseit.

Az első csoport entitásai a következők:

- *Mű (Work)*.
- *Kifejezési forma (Expression)*.
- *Megjelenési forma (Manifestation)*.
- *Példány (Item)*.

Közülük a „Mű” kizárólag absztrakt entitásként értelmezhető, fizikai megjelenése önmagában nincs, a „Kifejezési forma” még mindig absztrakt szint, az entitásnak fizikai jellemzője nincs. A „Megjelenési forma” és a „Példány” azok a tárgyasultság szintjén is értelmezhető entitások, amelyek fizikai jellemzőkkel is bírnak. Ezek közül természetesen csak a „Példány” az, ami fizikai valójában is megjelenik a gyűjteményben, a „Megjelenési forma” a „Példány” felől tekintve csak, mint a közös tulajdonságokkal bíró példányok halmazának leírása, nyomtatott művek, pl. könyvek esetében a kiadás fogalmával megfeleltethető. Ez az a pont, ahol fontos megjegyeznünk – még ha e megjegyzés triviálisnak is tűnik – hogy az FRBR modell alapvetően a fizikai valóságon, azon a hagyományos könyvtári modellen alapszik, amelyben a „Példány” létrehozójától (pl. kiadó és nyomda) bekerül a gyűjteménybe és a gyűjtemény általi legjellemzőbb szolgáltatás ennek a konkrét fizikai példánynak a továbbadása (kölcsonnése) a használó felé. A digitális gyűjtemény alapvetően másként működik: ott a használó hozzáférést kap a tárolt példányhoz és amennyiben letölti, akkor fizikailag is új példány keletkezik. Ennek az újfajta példányosításnak messze ható következményei vannak mind az objektumok azonosítása, mind a hozzájuk kapcsolódó jogok kezelése kapcsán! A „Mű (Work)” önálló szellemi vagy művészi alkotás. „A mű elvont entitás; nincs egyetlen olyan anyagi tárgy, amelyre valaki műként mutathatna rá. A művet egyedi megvalósításokon vagy kifejezési formákon keresztül ismerjük, de maga a mű csak a különböző kifejezési formák közötti tartalmi hasonlóságban létezik.” (FRBR HUN 24. p.) Az entitás fontosabb ismérvei a következők: van

- címe,
- típusa,
- keletkezésének ideje,
- célközönség,
- környezet.

A „Kifejezési forma (Expression)” „az a sajátos szellemi vagy művészi forma, amelyet a mű minden egyes alkalommal felvesz, amikor megvalósítják. A kifeje-

zési forma például sajátos szavakat, mondatokat, bekezdéseket stb. ölel fel, amely a mű megvalósításának eredménye szöveges formában, vagy sajátos lejegyzést, frazeálást stb. egy zenei mű megvalósításának esetében. A kifejezési formának, mint entitásnak a határai meg vannak határozva, mindazonáltal a fizikai forma szempontjai, mint amilyen a betűkép és az oldaltördelés, ki vannak zárva, mint amelyek nem lényegesek magának a műnek a szellemi vagy művészi megvalósításában.” (FRBR HUN 26. p.)

Az entitás fontosabb ismérvei a következők:

- címe,
- műfaja,
- keletkezésének ideje,
- nyelve,
- terjedelme,
- környezete,
- kritikai fogadtatása.

A „Megjelenési forma (Manifestation)” a mű kifejezési formájának valamely fizikai megtestesülése. „A megjelenési formaként meghatározott entitás a dokumentumtípusok széles skáláját foglalja magában; beletartoznak a kéziratok, könyvek, időszaki (folytatódó) kiadványok, térképek, plakátok, hangfelvételek, filmek, videók, CD-ROM-ok, multimédiák stb. A megjelenési forma, mint entitás valamennyi fizikai tárgyat képviseli, amelynek azonosak a jellemzői mind a szellemi tartalom, mind a fizikai forma tekintetében.” (FRBR HUN 28.p.)

Az entitás fontosabb ismérvei a következők:

- címe,
- szerzőségi közlés,
- kiadásjelzés,
- megjelenés helye, ideje,
- kiadó,
- fizikai hordozó,
- hordozó terjedelme,

a megjelenési forma azonosítója,

- betűtípus, betűméret,
- lapszámozás,
- fájljellemezők (elektronikus információforrás esetében),
- a hozzáférés módja (távoli hozzáférésű elektronikus információforrás esetében),
- hozzáférési cím (távoli hozzáférésű elektronikus információforrás esetében).

A „Példány (Item)” a megjelenési forma egy darabja. „A példányként meghatározott entitás egy valószínű darab. Sok esetben egyetlen fizikai tárgy (pl.

egy egykötetes monográfia egy példánya, egyetlen hangszalag stb.). Bizonyos esetekben a példányként meghatározott entitás egynél több fizikai tárgyat foglal magában (pl. két kötetben kiadott monográfia, három különálló CD-n megjelent felvétel stb.). A szellemi tartalom és a fizikai forma szempontjából a példány a megjelenési formában testesül meg, azaz általában azonos magával a megjelenési formával. Mégis előfordulhat, hogy az egyik példány különbözik a másiktól, még akkor is, ha a példány ugyanazt a megjelenési formát szemlélteti. Ezekben az esetekben a változatok a megjelenési formát előállító szándékától eltérő külső események következtében jöttek létre (pl. a kiadott példány a megjelenést követően sérült meg, a könyvtári kötetés következtében jött létre stb.)” (FRBR HUN 30–31.p.).

Az entitás fontosabb ismérvei a következők:

- példányazonosító,
- ujjlenyomat (fingerprint),
- a példány eredete (provenienciája),
- jelölések/ajánlások (dedikációk),
- kiállítás-történet,
- a példány állapota,
- kezelési történet,
- előírányzott kezelés,
- a példány hozzáférési korlátai.

A kiemelt részek mutatják, hogy a modell alkotói csak a fizikai tulajdonságokkal is rendelkező absztrakciós szinteken, tehát a „*Megjelenési forma*” és a „*Példány*” szintjén tartották szükségesnek azonosító szerepeltetését. A megjelenési forma azonosítójának célja, hogy a megjelenési formát minden más megjelenési formától megkülönböztesse. Az azonosító származhat központi (pl. nemzetközi) forrásból, mint pl. az ISBN, de akár helyi, csak a gyűjteményen belül hatókörrel bíró forrásból is. Egy megjelenési formához több azonosító is kapcsolható (FRBR HUN 50.p.). Ennek az azonosítónak a célja tehát csak a megkülönböztetés és az egyediség biztosítása, de nem célja pl. a fellelhetőség támogatása. Azonban, mint a „*Példány*” fölérendelt fogalmának azonosítója, nyilván csoportazonosító szerepe is van.

A modell a „hozzáférési cím” is a megjelenési forma adatai között szerepelteti, távoli hozzáférést elektronikus információforrások esetében. Példaként csak az URL-t nevezi meg. Kérdéses és vitatható, hogy az URL valóban a megjelenési forma adata-e, vagy a példányé? Amennyiben azt feltételezzük, hogy egy digitális gyűjteményben csak egy digitális dokumentumra igaz az, hogy rendelkezik mindazokkal a

kiadásra jellemző adatokkal, amelyek megkülönböztetik más kiadásoktól (pl. formátum), akkor igen. De ha pl. ugyanazon fájl több helyen, több URL-en is elérhető, akkor már kétségbe vonható. Érthető azonban a modellalkotó szándéka, miszerint a használó célja megtalálni egy elektronikus információforrást és hozzáférni ahhoz. Kevésbé valószínű, hogy bármely általánosnak tekinthető felhasználói szándék arra irányulna, hogy pl. a különböző URL-eken fellelhető, egyébként minden jellemzőjében megegyező fájlok közül megkülönböztetést tegyünk a bibliográfiai adatok szintjén. A polémia azonban jelzi, hogy a digitális objektumok esetében az absztrakciós szintek határai eltérőek lehetnek a hagyományos dokumentumok világában megszokottaktól.

A „példányazonosító” a modell szerint egyértelműen az egy gyűjteményen belüli megkülönböztető jel, ami a gyűjteményen belül a példányok megkülönböztetésére szolgál. A „megjelenési forma azonosítója” tehát lehet olyan jelzet, amely gyűjteményeken átívelő (pl. ISBN) és megkülönböztető funkciója mellett azonosító funkció is kapcsolódik hozzá, azaz segíthet megtalálni a különböző gyűjteményekben található, de ugyanazon megjelenési formához tartozó példányokat, a példányazonosítónak azonban csak megkülönböztető és általában hozzáférést biztosító (pl. raktári jelzet) funkciója van.

Érdekes még az ujjlenyomat (fingerprint) szerepeltetése példányadatként. Az FRBR modell elsősorban régi könyvek azonosító jegyeként tekint rá:

„Az ujjlenyomat (fingerprint) a kiadvány szövegéből vett, az egy adott nyomtatású kiadás azonosítására szolgáló betűjelek csoportja. A technikát elsősorban régi könyvek egyedi példányainak megkülönböztetésére használják. Különböző eljárásokat alkalmaznak az ujjlenyomat alkotására (pl. az *Institut de Recherche et d'Histoire des Textes in Paris* kiadásában megjelent *Fingerprints = Empreintes = Imprime leírás szerint*).”

Először is a megfogalmazásból nem világos, hogy a megjelenési forma (kiadás) adatáról vagy ténylegesen példányadatról van-e szó? „Az egy adott nyomtatású kiadás” megfogalmazás sokkal inkább arra enged következtetni, hogy pl. ősnymtatványok esetében a kiadásjelzés hiányában e jelek alapján lehet „kiadáshoz” rendelni egyes példányokat. Ennek alapján inkább a megjelenési forma ismérvei között lenne a helye. Fontos mindenesetre, hogy a modell alkotói figyelembe veszik azt, hogy a dokumentumoknak – mindegy, hogy megjelenési forma vagy példány szintjén – de lehetnek olyan természetes azonosító

jegyei, amelyek felhasználói relevanciával bírnak a keresés és használat során.

Azonosító rendszerek

A továbbiakban azokkal az azonosító rendszerekkel foglalkozunk, amelyek a jelenlegi gyakorlat szerint elterjedtek mondhatók digitális gyűjteményekben. Mivel a digitális könyvtár egyik elfogadott definíciója szerint az interneten keresztül elérhető információs rendszer (ARL, 1995), ezért a vizsgálatot azokkal az azonosítókkal kezdjük, amelyek az interneten általánosnak tekinthetők. Ezek az azonosítók a különböző internetes kiszolgálókon elhelyezkedő források elérhetőségét (címezését és azonosítását) szolgálják. Látni fogjuk, hogy ezek az azonosítók olyan elveket és technikákat valósítanak meg, amelyek természetesen a digitális könyvtárak számára sem megkerülhetők, sőt, a bemutatásra kerülő, speciálisan digitális könyvtári típusú források és gyűjtemények számára kialakított azonosítórendszerek ugyanezen elvek mellett és technikák alkalmazásával működnek.

Azonosító rendszerek az interneten

A források fellelhetőségének biztosítása az internet megjelenésével egyidejű probléma. A feladat az, hogy távoli gépeken, az arra a kiszolgálóra jellemző fájlstruktúrában hogyan lehet megtalálni, és mindenki számára elérhetővé tenni a forrásokat. Az elérhetővé tétel problémája címzési probléma. Tulajdonképpen nem áll másból, mint a távoli gép egyértelmű azonosításából (=egyedi névvel való ellátásából) és a belső elérési út leírásából. A kiszolgálók egyedi azonosításának eszköze az interneten az IP-cím, melynek kiosztása ellenőrzött módon és központosítottan történik. A nyilvántartásért az Internet Assigned Numbers Authority (iana.org) felel, regisztrálni pedig regionális (országos) regisztrációs pontokon lehet, az egyediség egyértelmű és teljes. Az IP-cím olyan értelemben beszélő kód, hogy az egyes számcsoportok (IPv4 esetén 4, IPv6 esetén 8) hozzárendelhetők ahhoz a regisztrációs központhoz és azon belül hálózathoz, alhálózathoz, ahonnan a kiosztás megtörtént. Mivel azonban az IP-cím mnemotechnikailag eléggé nehezen kezelhető, plusz egyes kiszolgálók fizikailag akár több IP-címen is jelen lehetnek, a World Wide Web kialakítása során Tim Berners-Lee új, az IP-re alapuló, de azt a felhasználó elől elfedő, könnyen megjegyezhető struktúrát vezetett be; ez az URL. Szintén az internettel egyidejű probléma, hogy sem

az IP-cím, sem az elérési út nem mond semmit a forrás tartalmáról, a felhasználó pedig már csak olyan esendő, hogy nem a technikai címekre, hanem a tartalmakra kíváncsi. Meg kellett oldani tehát, hogy olyan helyek is legyenek, ahol tartalmi jellemzők alapján azonosíthatók a források. Ez a fajta azonosítás azonban már nem feltétlenül az egyediséget, hanem a csoporthoz tartozást hivatott szolgálni. A csoporthoz tartozás fogalmát azonban több szinten is lehet értelmezni – mint azt a könyvtártudomány néhány évezred óta gyakorolja. Vannak források, amelyek néhány fontos vagy kevésbé fontos jellemzőjükben egyeznek meg, és vannak olyanok, amelyek mindenben (a „minden” fogalmát, most tekintsük általánosnak, mondjuk, legyen igaz két fájlra, amelyek egymás másolatai, de az interneten különböző helyeken lelhetők fel). Ezt a rendkívül erős csoportképző funkciót szintén Tim Berners-Lee olyan fontosnak érezte (joggal), hogy kezelésére külön azonosítót vezetett be; ez az URN.

URL – Uniform Resource Locator

A legismertebb és legelterjedtebb azonosító az interneten az URL. Mint a neve is jelzi, ez az azonosító helymegjelölésre szolgál, a népszerű neve – webcím – is ezt jelzi. Az URL alapvetően négy elemből áll, amelyek mind a kívánt forrás fellelhetőségét biztosítják:

- a kommunikációs protokoll típusa (pl. http, ftp, mailto stb.),
- a gép (szerver vagy tartomány) neve vagy azonosítója, amelyen a forrás megtalálható (pl. www.oszk.hu),
- a hálózati port számát, amin a forrás elérhető (pl. :430) – gyakran alapértelmezett a port, tehát elhagyható,
- a fájlhoz vezető elérési utat a célgépen belül (/ {könyvtárnév} / {fájlnév}).

Pontos struktúráját és a felhasználás szabályait az IETF RFC 1738 szabvány írja le: <https://tools.ietf.org/html/rfc1738>.

Az URL használata egyszerű, a legszélesebb körben elterjedt – ugyanakkor mégsem problémamentes. Addig rendben van, amíg a fájlok az URL-ben feltüntetett helyen valóban megtalálhatók, de ha például akár csak a könyvtárstruktúra megváltoztatása miatt, megváltozik az elérési út, az URL máris pontatlan. Mint ahogy attól sem véd meg semmi, ha az adott címen megváltozik a tartalom. Ez esetben a használó

egészen más forráshoz fér hozzá, mint szeretett volna. Jó esetben semmihez, vagy nyilvánvalóan téveshez, de azt sem zárja ki semmi, hogy szándékosan hamisítotthoz. Mégis, mivel az URL egyértelműen a forrás fizikai címére utal – könyvtári szóhasználattal mondhatnánk, annak „raktári jelzete” (bár akkor az FRBR modellben is példányadatnak kellene lennie) – a digitális források címhez kötött azonosításában megkerülhetetlen.

URN – Uniform Resource Name

A felismerés, hogy az URL korlátozottan használható, elég nyilvánvaló volt ahhoz, hogy már Tim Berners-Lee olyan megoldást keressen, amely lehetővé teszi, hogy a webes források egyéb megkülönböztető jegyeit is kezelni képes azonosítót alkosson. A legismertebb egyedi forrásazonosító az URL mellett az URN. Az URN nem a helyet azonosítja, hanem magát a forrást, méghozzá a neve alapján. Meg szokták jegyezni, hogy a „forrás” jelen esetben nem feltétlenül kell, hogy internetes forrás legyen, lehetnek akár elvont dolgok nevei is. A név akkor lehet azonosító elem, ha egyedi. Az egyediséget pedig mindig csak egy adott értelmezési tartományon belül lehet értelmezni. Az URN felépítése során tehát az egyedi név (amely az URN utolsó eleme) mellett meg kell adni azt az ún. névteret is, amin belül az adott név egyedi. Ilyen névtér lehet pl. az ISBN, amelyen belül a számok, képzésük szabályánál fogva egyediek. Egy ISBN számot, mint egyedi nevet tartalmazó URN felépítése tehát a következő: urn:isbn:0451450523. A bevezető urn karaktersor az URN sémát jelöli. Ez a séma szintaktikai szerepét tekintve megfelel az URL-ben szereplő kommunikációs protokoll típusának. A különbség annyi, hogy míg a kommunikációs protokoll típus hatására a böngésző végrehajt egy utasítást (pl. mailto hatására megnyitja a default levelező rendszert), addig az urn: hatására nincs mit végrehajtania. Ahhoz, hogy egy urn sémajelölést követő névtér: egyedi_név karaktersorozatot értelmezhetővé tegyünk, vagy találjunk a neten olyan forrást, amelyre igaz, hogy az egyedi név jellemezi, ún. URN feloldóra van szükség. Ennek pedig tipikusan része egy url, amely a feloldóhoz vezet, és amely url-en belül az urn paraméterként a konkrét egyedi névvel jelzett forrást jelöli.

URI – Uniform Resource Identifier

Úgy az URL, mint az URN az URI-k családjába tartozik. Tim Berners-Lee 1994-ben még Universal

Resource Identifier-ként nevezte el az azonosítót és mint az URL és URN fölérendelt fogalmát, e két azonosító rendszer közös szabályrendszerét írta le (RFC 1630, Universal Resource Identifiers in WWW. A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web). Később változott meg a neve Uniform Resource Identifierre és a két kategória (URL és URN) helyett számos séma által leírt különböző, de szintaktusában hasonló felépítésű azonosító fölérendelt fogalma lett. Aktuális állapotát az RFC 3986 Uniform Resource Identifier (URI): Generic Syntax szabályrendszer írja le.

Az URI általános felépítése a

„scheme (séma) – authority (hatókör vagy értelmezési tartomány) – path (elérési út) – query (lekérdezés) – fragment (darab vagy rész)” struktúrának felel meg. *Séma* az URL-ek http:, ftp:, mailto:, stb. előtagja és az urn: megjelölés.

Az *authority* körébe tartozik az URL gépneve vagy az URN névtér azonosítója.

A *path* minden URI kötelező eleme a sémával együtt, a *query* és a *fragment* nem kötelező elemek.

A sémák kiadásáért a domain regisztrációért is felelős Internet Assigned Numbers Authority (IANA) (www.iana.org) felel, jelenleg 279 sémát tartanak nyilván.

Digitális gyűjtemények speciális azonosító rendszerei

Az alábbi néhány rendszert példaként emeljük ki, nem célunk minden digitális könyvtári azonosítórendszer bemutatása. Látni fogjuk, hogy a legelterjedtebb azonosítók felépítése, képzési módja és technikai háttere nagyon hasonló: mindegyik az URN – URL megfeleltetés megvalósításán alapul, tulajdonképpen mindegyik egy URI.

A Handle rendszer

A Corporation for National Research Initiatives (CNRI) már a 90-es években együtt dolgozott a kongresszusi Könyvtárral a National Digital Library Projecten (<http://www.cnri.reston.va.us/home/cstr.html>). E projekt keretén belül kezdték meg egy keretrendszer kidolgozását, amely – azon túl, hogy a digitális könyvtárak számára példaként szolgált – olyan megoldások forrása lett, amelyek azután széles körben elterjedtek, mint más digitális gyűjtemények által igénybe vett szolgáltatások. (Kahn – Wilensky, 1995.). A keretrendszer egyébként a „digitális objekt-

tumra” vonatkozólag is ad egy meglehetősen technikai definíciót (Arms, 1997.), de ennél sokkal lényesebb eleme a „handle” azonosítók bevezetése. A „handle” a CNRI megfogalmazása szerint: olyan általános célú azonosító („general purpose identifier”), amely alkalmas az interneten elérhető források, pl. digitális objektumok azonosítására hosszú távon, függetlenül attól, hogy milyen repositoriban vagy adatbázisban tárolják őket. A „handle” azonosító hosszú távon fennmaradó név, de a forrás, amit azonosít, közben megváltoztathatja a formáját, tárolási helyét, vagy más módon is megváltozhat (!) (Arms, 1997). A CNRI által a „handle” azonosítók kezelésére kifejlesztett Handle rendszer egy olyan osztott számítógépes rendszer, amely tárolja a „handle” azonosítókat és a hozzájuk tartozó adatokat, amelyek alapján az azonosított objektumok fellelhetőek és elérhetőek. (Arms, 1997) Technikai megvalósítását tekintve a Handle rendszer tehát nem más, mint egy URI feloldó. Emellett a Handle rendszer szolgáltatásként is működik más szereplők, így pl. a DOI rendszer üzemeltetői számára. Ezek a szereplők Handle azonosítókat vásárolhatnak, pontosabban egy tartományt, amin belül aztán saját hatáskörben adhatnak ki további azonosítókat. Ezek mindegyike az elsődleges szereplő (pl. DOI) által megvásárolt tartományt jelző prefixszel kezdődik. A prefixek kiadását 2015-ig a CNRI végezte, azóta ezt a jogot átadta a svájci székhelyű DONA non-profit alapítványnak. (https://www.dona.net/handle_system/).

DOI – Digital Object Identifier

Formailag a DOI (Digital Object Identifier = digitális objektum azonosító) egy URI névtér az info: sémán belül. A DOI szabványos leírását az ISO 26324:2012 szabvány tartalmazza. A rendszer kialakításának célja elsősorban az interneten fellelhető tudományos publikációk (folyóiratcikkek, disszertációk stb.) folyamatos fellelhetőségének biztosítása volt. Szükségességét ugyanúgy az URL korlátaival indokolják, mint minden URI-jét. Ennek megfelelően a DOI azonosító egy objektumhoz permanensen hozzárendelt azonosító, amelynek feladata tartós kapcsolat biztosítása az objektumról való információk eléréséhez az interneten keresztül. Az információk körébe beletartozik az objektum helye, de a róla szóló információk, meta-adatok elérhetősége is. A DOI azonosító feloldásáért a DOI rendszer felel, azaz a DOI rendszerben vannak nyilvántartva az adatok elérhetőségei, amelyek az azonosítóval ellátott objektumhoz kapcsolódnak, így

pl. az objektum URL-je. Amennyiben egy objektum több URL-en is elérhető, úgy akár az összes URL is fellelhető a DOI rendszerben – feltéve, ha annak bejelentése megtörtént.

A DOI azonosító szintaxisa: 10.{SZERVEZET AZONOSÍTÓJA}/{EGYEDI AZONOSÍTÓ}.

A 10. prefix a DOI által a DONA-tól vásárolt Handle azonosító, a „szervezet azonosítóját” a DOI regisztrációs központokon keresztül lehet igényelni, az egyedi azonosítót pedig a regisztrált szervezet adja ki, úgy, hogy a szervezetazonosítón belül egyedi legyen. A képzés módját tekintve nem lehet nem észre venni a hasonlóságot az ISBN azonosítók kiosztásával. A DOI népszerű neve éppen ezért a „digitális objektumok ISBN-je”. Ráerősít erre a névre az a tény is, hogy a DOI-ban regisztrált szervezetek közül meglehetősen sok a kiadó. A DOI azonosító általános URI formátumban is hozzárendelhető egy objektumhoz (doi:10.1000/182), de lehetőség van arra is, hogy URL-ként rögzítsük (<https://doi.org/10.1000/182>). Ez utóbbi esetben a dx.doi.org szerver szólíttatik meg, amely DOI feloldóként elvezet az objektumhoz (jelen esetben a DOI kézikönyvhöz). A DOI, pontosabban a kiadásáért felelős International DOI Foundation a Handle rendszer infrastruktúráját használja.

ARK – Archival Resource Key

A California Digital Library által 2001-ben kifejlesztett rendszer elsősorban digitális archívumok kezelőinek nyújt segítséget. Az azonosító bármilyen objektumtípusra használható, beleértve a digitális objektumokat, fizikai objektumokat, személyeket, személyek csoportjait, de még elvont objektumokat is (<https://confluence.ucop.edu/display/Curation/ARK>). Ennek az azonosítónak sem más a célja, mint az, hogy a „névvel”, ami egyébként egyedi, de nem biztos, hogy beszélő kód, egyértelműen azonosíthatók legyenek objektumok, amelyek adott esetben eltérő URL-en érhetőek el. Ennek megfelelően az ARK azonosító szintaktikailag az URL-ből vezethető le:

<http://example.org/ark:/12025/654xz321/s3/f8.05v.tiff>, ahol az ark: prefix előtti rész egy ún. „Name Mapping Authority Hotspot” (NMAH), amely egy olyan kiszolgáló domain, amely alkalmas a mögötte álló kifejezés feloldására. Ez azt jelenti, hogy különböző NMAH-k ugyanarra az ark:-kal kezdődő azonosítóra ugyanazt az eredményt kell, hogy adják. Az ark: label után következik a „Name Assign Authority Number” (NAAN), ami annak a szervezetnek az azonosítója, amelyik elsőként azonosította az objek-

tumot. Jelenleg több mint 470 szervezet szerepel a szolgáltatás honlapján, mint NAA, közöttük olyanok, mint a California Digital Library, az Internet Archive, a British Library vagy a Google. Az ötjegyű NAAN utáni rész a név, és olyan minősítők, amelyeket az NMA támogat (Kunze, 2013). Látható, hogy az ARK is a különböző helyen fellelhető, ugyanazzal az azonosítóval jellemzett, de nem feltétlenül egymásnak mindenben megfelelő objektumok elérését célozza.

Példányosítás a digitális könyvtárban

Ha a fent bemutatott rendszereket összevetjük a DLRM-mel és az FRBR modellel, akkor azt kell látnunk, hogy ezek az információs objektum entitás különböző absztrakciós szintjein alkalmazhatóak. Az URL egyértelműen a példány szintjén, a DOI a megjelenési forma szintjén, az URN pedig szintén elsősorban a megjelenési forma szintjén. Okozhat némi zavart, hogy a DOI és az URN esetében nem ritka, hogy a feloldók csak egy URL-t tartalmaznak, azaz úgy tűnhet, mintha a megjelenési forma csak egy példányban létezne, de ez nem változtat azon, hogy minden példány, amely ugyanazon DOI-val vagy URN-nel jellemezhető, ugyanazon megjelenési formához tartozik. Könnyű tehát belátni, hogy bármikor előkerülhet az interneten olyan példány, ami ennek a kritériumnak megfelel, tehát megkaphatja ugyanazt a DOI-t vagy URN-t.

De hogyan hat ez egyetlen gyűjteményre? Ha az URL-t szimplán a könyvtárban megszokott raktári jelzettel azonosítjuk, akkor valószínűleg sehogy. A digitális gyűjtemény egyik sajátossága a hagyománnyal szemben éppen az, hogy a szolgáltatás nincs a fizikai példányok számához kötve. Az FRBR modell „obtain (megszerez)” felhasználói funkciója is megengedő, amennyiben digitálisan tárolt, online elérhető információforrások esetében kielégítőnek tartja a hozzáférés biztosítását is, a letöltés már nem kötelező elem (FRBR LRM 10.p.). Ebből adódóan a digitális könyvtár működhethet akár megjelenési formánként egy példányos gyűjteményként (tekintsünk most el a biztonsági másolatoktól, és az „egy példány” fogalmát korlátozzuk a szolgáltatási példányra!), akkor pedig pl. a DOI akár elégséges azonosító is lehetne. Számos érv szól azonban emellett, hogy ez a felfogás nem kielégítő.

1. Az első, triviális érv, hogy egy objektum letöltésével mindenképpen új példány keletkezik.

Igaz rá, hogy az URL-t kivéve, minden tartalmi és formai jellemzőjében megegyezik a szolgáltatási példánnyal. Mondhatná a könyvtár, hogy jó, de az a példány már kívül esik az ő hatókörén, nem felelős sem a megőrzéséért, sem a további felhasználásáért. Ez mind igaz, de egy felelőssége mindig meglesz: azt a példányt vissza kell tudni vezetni a szolgáltatási példányhoz. Vissza kell tudni vezetni, mert pl. a tudományos művek korrekt hivatkozása csak így valósítható meg. A könyvtár gyűjteményeiből szolgáltatott digitális tartalmak hitelességéért a könyvtár tartozik felelősséggel, ugyanúgy, mint ahogy a hagyományos, analóg kiadványok esetében a kiadó. Ez a felelősség több szintű: alapesetben arra terjed ki, hogy a könyvtár gyűjteményéből letöltött anyag megegyezik az oda hiteles forrásból feltöltöttel. Ez lényegében annak felel meg, hogy az analóg világban egy kiadás minden példánya egyforma. Ez ott a kiadó (és a nyomda) felelőssége. A példányosítás a digitális világban viszont nem a feltöltéssel, hanem a letöltéssel valósul meg. Egy tudományos munka gondos szerzője például adott könyvtár gyűjteményében tekinti meg, és ha nagyon gondos és jogosult, akkor tölt le egy hivatkozni szánt művet. Számára a hitelesség forrása inentől a könyvtár és nem a kiadó. Ha ellentmondásba keveredik, mert valaki kétségbe vonja a hivatkozása helyességét, akkor a könyvtárra fog mutogatni és az onnan letöltött példányt fogja lobogtatni a digitális szélben. A könyvtár felelőssége annak bizonyítása, hogy a letöltött példány hiteles és megegyezik az adott DOI azonosító alatt, más gyűjteményekben elérhető példányokkal.

2. Vannak olyan helyzetek, pl. jogvédett tartalmak példányosítása esetén, amikor azt is pontosan tudni kell, hogy ki, milyen jogosultság alapján jelentette meg (!), vagy töltötte le a példányt. Ebben az esetben tehát nemcsak a példány és az eredeti egyezőségét kell tudni bizonyítani, hanem a létrejött példányhoz új adatokat is kell tudni rendelni. Ilyen adat pl. a jogosítottság, a jogosultság ellenőrzésének eredménye, az (új) tulajdonos adatai. Nem véletlen, hogy az FRBR Library Reference Modell is kiemeli az „Agent” entitás (amelynek megjelenési formája lehet a „User” is) tartalmi entításokhoz kötődő kapcsolatai között az „Item” – „Agent” viszonylatban az „owns” kapcsolat lehetőségét. (Lásd az ábrákat!)

- rán, másolással (file copy) keletkezett,
- tulajdonosa nem a könyvtár, hanem a felhasználó, de az ezzel kapcsolatos meta-adat adott esetben a könyvtárban, a szolgáltatási példányval összefüggésben megőrzendő,
- a felhasználás restriktiókkal is járhat, amelyre vonatkozó meta-adatok szintén a könyvtár rendszerében megőrzendők lehetnek.

Elsősorban a két utóbbi ismerv miatt belátható, hogy a felhasználói példányt a szolgáltatási példánytól meg kell különböztetni, azaz olyan azonosítóval kell ellátni, ami elsősorban (globálisan egyedi) megkülönböztető funkcióval bír. Azt is beláthatjuk, hogy a könyvtár számára is fontos, hogy a gyűjteményében fellelhető (szolgáltatási) példányokat megkülönböztessük az esetlegesen más gyűjteményekben fellelhető példányoktól, hiszen a könyvtár joga és felelőssége mindig csak a saját gyűjteményével és az abból letöltött példányokkal kapcsolatban áll fenn.

Követelmények az egyedi példányazonosítókkal szemben

A fenti okfejtést elfogadva megállapíthatjuk, hogy a kutatás elején feltett 4. kérdésre, miszerint: „Lefedik-e ezek az azonosítók a digitális gyűjtemények minden igényét?” a válasz, hipotézisünknek megfelelően nemleges. Jelenleg azonban még csak az világos, hogy van számos olyan pont, ahol a jelenlegi rendszerek nem használhatók. Ahhoz, hogy egy hitelesítő rendszer prototípusát kidolgozzuk, szükség van egy követelményrendszerre, ami azt mondja meg, minek kell megfelelni. Az alábbiakban ezeket a követelményeket szedjük össze és teszünk kísérletet rendszerbe foglalásukra.

Egyediség

Definiáljuk a példányt, mint a digitális gyűjtemény legkisebb, tovább már nem bontandó egységét. Tekintsünk most el attól, hogy összetett dokumentumok esetén a legkisebb egységek („item”) akár fizikailag is más-más helyen tárolt elemek lehetnek, amelyek bibliográfiai értelemben mégis együttesen alkotják a példány fogalmát. A példány fogalmához hozzá tartozó definíció még, hogy az FRBR modell szerint „szellemi és művészi tartalom, valamint formai jegyek tekintetében a megjelenési formát példázza, gyakran [e jegyek tekintetében] azzal megegyező.” (FRBR LRM 18.p.)

A definícióból nyilvánvaló tehát, hogy a példányazonosítónak megkülönböztető funkcióval kell bírnia

(értsd: nem elég az egyértelmű visszakereshetőség funkció) az egyébként minden egyéb elemében megegyező példányok esetében is. Az azonosítást meg lehet oldani beszélő kódokkal, vagy nem beszélő kódokkal. A beszélő kódok használata prekoordinált kód kiosztást jelent és mindenképp csoportosító jellegű. Csoportosító jellegéből fakad, hogy az egyediség értelmezési tartománya csak a csoportképzésen belül értelmezhető, azaz, hogy feltételezi, hogy az azonosítandó példányok közötti különbségtétel csak is és kizárólag a csoporthoz tartozás viszonylatában értelmezhető. Erre lehetne példa, ha mondjuk a DOI azonosítót próbálnánk meg alkalmazni példányazonosítóként. Ki tudnánk fejteni vele, hogy egy-egy azonosított példány ugyanannak a megjelenési formának a példázata, de például egy, a szerző csoportképző tulajdonságát szem előtt tartó különbségtételt már nem. Ráadásul a kiosztás módja is problémás lenne, hiszen a példányosítás nem a kiadó hatáskörében, hanem az egyes könyvtárakéban történik, tehát vagy minden könyvtárnak minden esetben a kiadóval egyeztetve kellene új DOI azonosítót kiadni, vagy az ütközések elkerülhetetlenek. És akkor még nem beszélünk arról a kérdésről, hogy ki a DOI azonosító tulajdonosa és ki a költség viselője? Akkor sem járunk sokkal jobban, ha a DOI prefixet a könyvtár igényli és nem a kiadó, hiszen így pont a DOI csoportképző ereje (kiadásazonosítás) veszik el. A DOI jelen esetben csak egy példa, az okfejtés minden prefixszel operáló, csoportképzésen alapuló beszélő kódra igaz.

Igaz továbbá, hogy ha a beszélő kód szemantikai szinten nem értelmezhető a legutolsó jegyig (mint mondjuk egy ETO szám), akkor törvényszerűen illetleni kell hozzá egy csak szintaktikai szinten értelmezhető – tehát csak megkülönböztető jelleggel bíró – nem beszélő kódot. A példány fenti definíciójából következik, hogy az ilyen elem megléte elengedhetetlen példányazonosítóknál. Ezt általában és legegyszerűbben sorszámok kiosztásával szokták megvalósítani. Mivel azonban a példányosítás során a példány egyrészt kikerül a könyvtár hatóköréből, másrészt mindig biztosítani kell, hogy ennek ellenére is visszavezethető legyen a könyvtárra és az ott tárolt szolgáltatási példányra, az egyediség értelmezési tartományát sokkal tágabban kell értelmezni, mint a könyvtár gyűjteménye.

„Ex Libris”, avagy hitelesítő funkció

A fentiekben bizonyítottuk, hogy a könyvtár felelőssége szempontjából követelmény, hogy az ál-

lományából letöltéssel keletkező példány mindig visszavezethető legyen a szolgáltatási példányra. A visszavezethetőség során bizonyítani kell, hogy a letöltéssel keletkezett példány:

- valóban a szolgáltatási példány másolataként jött létre,
- minden lényeges elemében (elsődlegesen tartalmában) megegyezik az eredeti példánnyal,
- ebből következően letöltése óta nem változtatták meg.

Első lépésként tehát meg kell találni az eredeti példányt. Az egyedi azonosítókat tehát úgy kell kiosztani, hogy az eredeti példány egyedi azonosítója és a másolat egyedi azonosítója egymásnak megfeleltethető legyen. Ez két módon érhető el. Vagy úgy, hogy az eredeti példány azonosítójából képezzük a másolati példány azonosítóját, mondjuk egy prefix. suffix struktúrával, vagy pedig az URN feloldókhoz hasonló megfeleltető szolgáltatással biztosítjuk a kapcsolatot. Miután az azonosító és az URL közötti feloldó funkcióra mindenképpen szükség van, nem szükséges ragaszkodni a származtatott azonosító képzéséhez.

Következő lépésben vizsgálni kell, hogy a másolati példány megegyezik-e az eredetivel? Ennek a technika jelen állása mellett leghatékonyabb eszköze digitális dokumentumok esetében a kriptográfiai hash függvény. E függvény – számos egyéb tulajdonsága mellett, amelyek pl. az azonosítók képzésének során is kihasználhatók – rendelkezik azzal a jó tulajdonsággal, hogy tetszőleges hosszúságú input (jelen esetben digitális dokumentum) esetében akkor és csak akkor adja ugyanazt – az esetünkben az inputnál jóval rövidebb – eredményt, ha a két input bitről bitre megegyezik egymással. Abban az esetben tehát, ha felmerül a kétség, hogy a másolati példány nem egyezik meg az eredetivel, úgy azt kell vizsgálni, hogy mindkettőre ugyanazt a kriptográfiai hash függvényt alkalmazva, ugyanazt a lenyomatot kapjuk-e. Ha a két lenyomat megegyezik, úgy a két input is – nagyon nagy valószínűséggel – megegyezőnek tekinthető. Csak az érdekesség kedvéért jegyezzük meg, a kriptográfiai hash függvénnyel létrehozott lenyomatot szokás a digitális tartalom ujjlenyomatának („fingerprint”) is nevezni. Ha a kifejezést összevetjük az FRBR példányadatok jellemzői között emlegetett „ujjlenyomattal”, mint azonosító funkcióval bíró adattal, akkor láthatjuk, hogy itt is hasonló funkcióról van szó: az ujjlenyomat megkülönbözteti egymástól a tartalmi és formai jegyeikben

eltérő példányokat egymástól, de az ilyen jegyeikben megegyezőket összeköti. Tehát az „ujjlenyomatként” van azonosító funkciója, de elsősorban a csoportosító jellegén keresztül, azaz az egymás másolataiként létrehozott példányok csoportját azonosítja és nem az egyes példányokat.

A könyvtár azonban nem csak szolgáltatója a digitális tartalmaknak, hanem előállítója, kutatója, minőségének javítója is. Ebbéli tevékenysége során pedig nem ritkán előfordul, hogy a digitális tartalmat, leggyakrabban annak szöveges részét megváltoztatja, annak érdekében, hogy filológiai értelemben magasabb minőségű tartalom szülessen. Ebben az esetben a probléma még összetettebb. Hiszen a módosítás előtt letöltött példány lenyomata nyilvánvalóan nem fog megegyezni a változtatáson átesett tartalommal. Sőt, ha a könyvtár gondosan akar eljárni, akkor a megváltoztatott tartalmat már nem is ugyanazzal az azonosítóval látja el, mint a változtatás előttit. A korábban letöltött példány hitelességét azonban mindig bizonyítani kell tudni, mint ahogy azt sem árt jelezni, hogy azóta létezik egy, a korábbinál magasabb minőséget képviselő tartalmi változat is. Mi történjen a korábbi változattal? Ez persze egyrészt gyűjteménypolitikai kérdés, de könnyen elképzelhető és indokolható, hogy a korábbi változatot a gyűjtemény nem kívánja megőrizni. Ebben az esetben úgy lehet biztosítani a visszavezethetőséget, ha a könyvtár megőrzi az eredeti azonosítót, és ahhoz kapcsolódóan a dokumentum lenyomatát, mint meta-adatot. Így megmarad az összehasonlítás lehetősége a korábban letöltött példányokra vonatkozóan is.

Végezetül emeljük ki, hogy a könyvtár a digitális térben egyfajta hiteleshelyként működik. Hiteleshelyként abban a tekintetben, hogy az ott fellelt tartalmak a használók számára megbízhatóbbak, tehát értékeesebbek, mint az interneten esetleg más forrásból megszerezhető tartalmak. A hitelesség forrása a könyvtár évezredek tradíciója. A könyvtár válogat; ami a gyűjteményébe bekerül, az – a gyűjtőkörü szempontokat figyelembe véve – értékeesebb annál, mint ami nem kerül be. Egy virtuális gyűjtemény esetében (virtuális gyűjtemény alatt értsük most az olyan információs objektumokból álló gyűjteményt, amelyek forrása az internet és a könyvtár csak hozzáférést biztosít hozzájuk) a felhasználó általában azzal a bizalommal fordul a könyvtárhoz, hogy az ott fellelhető dokumentumokat a könyvtárosok válogatták, tehát értékelték is. A könyvtárral szemben fennálló bizalom még olyan, egyre növekvő reputációval rendelkező forrásokkal szemben is fennáll, mint a Wikipedia (<https://>

en.wikipedia.org/wiki/Reliability_of_Wikipedia). Amennyiben a könyvtár meg akar felelni ennek a bizalomnak, sőt, esetleg profitálni is szeretne belőle, akkor az újabb érv amellet, hogy példányainak hitelességéért, azért a hitelességért, amit a könyvtár kölcsönöz számukra olyan azonosítót vezessen be, amely ezt a funkciót támogatja.

Letagadhatatlanság

Mivel a letöltött példánnyal kapcsolatban nem ritkán jogi (tulajdonjog, szerzői joggal kapcsolatos felhasználási korlátozás, jogosítottság stb.) információk is keletkeznek, a könyvtárnak úgy kell eljárnia, hogy a visszavezethetőségen túlmenően azt is bizonyítani tudja, hogy a másolati példány olyan körülmények között keletkezett, ami megfelel jogszabályi előírásoknak. Bizonyítani kell tudni például, hogy a szerzői jog védelme alá tartozó művekről csak úgy készülhet legális másolat, hogy azon elhelyezésre kerülnek a DRM-mel kapcsolatos jelzések. Ennek legegyszerűbb módja, ha maga a példányazonosító olyan, amin keresztül ez a kritérium biztosítható. Ez egyrészt megköveteli, hogy a példányazonosítók valóban a legtágabb értelmezési tartományon, térben és időben egyediek legyenek, másrészt, hogy a hozzájuk kapcsolódó jogi információk ne (csak) a dokumentumon, de a könyvtár hatókörében is tárolásra kerüljenek. A könyvtárnak tehát nem csak, hogy el kell látnia a keletkezett példányokat azonosítókkal, de ezekről az azonosítókról nyilvántartást is kell vezetnie, és a kapcsolódó példányadatokat meg is kell őriznie. Amennyiben pedig a könyvtár olyan azonosítórendszert használ, ami globálisan egyedi azonosítókat biztosít, úgy ennek az információnak – vagy legalábbis annak, hogy ezek az információk melyik gyűjtemény adatai között érhetők el – a szolgáltatónál is elérhetőnek kell lenniük.

Hamisíthatatlanság

Első olvasásra talán értelmetlennek tűnő követelmény, hogy egy könyvtári példányazonosító hamisíthatatlan is legyen, de a fentiek alapján talán már nem tűnik túlzásnak. A hitelesség és jogosítottság olyan súlyú adatai és információi kapcsolódhatnak egy példányhoz, a könyvtárra olyan jogok és felelőségek hárulnak, hogy azokat mindenképpen célszerű egy viszonylag egyszerű eszközzel védeni. A hamisíthatatlanság kritériuma egyébként szorosan kapcsolódik a globális egyediséghez: mindenképpen el kell kerülni, hogy bárki, akár véletlenül, akár a károkozás szándékával egynél többször kiadhassa

ugyanazt az azonosítót. Ebből következően először azt kell vizsgálnunk, hogy mi az azonosító képzési módja? A hamisítás elleni védelem ugyanis ott kezdődik, hogy senki se legyen képes olyan azonosító kód képzésére, amelyet a jogosított kódképző gyárt, és ne tudja azt illetéktelenül az azonosítórendszerbe helyezni. A számítástechnika számos olyan módszert ismer, amely segítségével nagyon nagy valószínűséggel egyedi, előre nem megjósolható azonosító gyártható. Ezek összefoglaló neve UUID (Universally Unique Identifier = univerzálisan egyedi azonosító). Közös jellemzőjük, hogy valamely előre nem, vagy csak nehezen jósolható inputból képzik az azonosítót, mint pl. idő, esetleg az idő és a kódképző gép azonosítója (mac address), esetleg véletlenszám. Mindezen képzési módoknak közös jellemzője, hogy az input tulajdonságaira építve nehezen, vagy egyáltalán nem jósolható, hogy mely azonosítók oszthatók ki az azonosító rendszerben. Jegyezzük meg, hogy egyébként a kriptovaluták, mint pl. a Bitcoin „bányászata” is azon az elven alapul, hogy előre nem jósolható kódokat kell képezni, és azok globális egyediségét biztosítani. Könyvtári példányok azonosítására azonban nyilván nem a Bitcoin, vagy bármely más drágán képzett kód a megfelelő eszköz, de mindenképpen olyan forrást kell választani, ami úgy a képzés, mint a felhasználás során gondoskodik az azonosító kódok hamisíthatatlanságáról.

Láthatóság, megjegyezhetőség

Mivel a digitális gyűjteményekből keletkező példányok száma nem, vagy csak nehezen becsülhetően sok, nyilvánvaló, hogy az azonosítók is csak nagyon hosszúak lehetnek, hiszen másképp nehezen biztosítható az egyediség. A hosszú azonosítók pedig az emberi szem és értelem számára nehezen megjegyezhetők. Mivel az azonosítók elsődleges funkciója nem is a fellelhetőség biztosítása, hanem a megkülönböztetés, ezért másodlagos fontosságú, hogy maga a kód az emberi szem számára látható módon van-e elhelyezve a dokumentumon. A gépi olvashatóság sokkal fontosabb. Digitális környezetben biztosan szükség van valamilyen, az URI-k esetében már ismerttetett feloldó használatára, tehát olyan szintaxist érdemes választani, ami annak elérését megkönnyíti. Ezzel együtt, mivel a kód hitelesítő funkcióval is bír, célszerű legalább annyit feltüntetni látható módon is a példányon, hogy azonosító kóddal el van látva. Ha az azonosító kódot nyomtatott példányon helyezzük el, akkor pedig célszerű QR-kód formájában rögzítenünk.

A kutatás eredménye és a Loca Credibilia könyvtári alkalmazása

A kutatás legfontosabb eredménye az egyediségre, hitelesítő funkcióra, letagadhatatlanságra és hamisíthatatlanságra vonatkozó kritériumrendszer, amit a digitális dokumentumok példányosítása, a felhasználói példányok létrejötte során az azonosítás során alkalmazni kell. A Loca Credibilia rendszerben ilyen azonosítót alkalmazunk, ez biztosítja a megbonthatatlan logikai kapcsolatot az azonosított dokumentumpéldány és a róla tárolt hitelesítő adatok között. Az azonosító képzése során ellenőrzöttek egyedi, csak egyszer felhasználható, nem beszélő kód. Nem helyettesít és nem vált ki más kódokat, célja annyi, hogy a biztonságos hitelesítést és a hitelesség ellenőrzését támogassa. Forrása a Loca Credibilia rendszer része, amelynek könyvtári alkalmazása az alábbiak szerint működik:

1. Minden, a könyvtár gyűjteményébe kerülő digitális dokumentumot egyedi, hamisíthatatlan és letagadhatatlan azonosítóval látunk el, mint azt a kutatás eredménye megkívánja.
2. Az azonosítással egy időben képezzük a dokumentum kriptográfiai hash függvényrel képzett digitális ujjlenyomatát.
3. Az azonosítót hozzárendeljük a dokumentumhoz, egyúttal az ujjlenyomattal és a képzés időpontjával együtt tároljuk a Loca Credibilia hitelesítő szerveren.
4. Amikor egy felhasználó, akár nevesített, akár anonim módon egy dokumentum letöltését vagy akár megjelenítését kezdeményezi, azaz megvalósul a példányosítás, az így létrejött egyedi, felhasználói példány új egyedi azonosítót kap, amelyet látható módon, egy „pecsét” formájában is elhelyezünk a dokumentumpéldányon. A pecsét hyperlink és QR-kód formájában is tartalmazza az egyedi azonosítót, valamint egy feliratot, ami igazolja, hogy az adott felhasználói példány a rögzített időpontban az Országos Széchényi Könyvtár gyűjteményéből került letöltésre. Opcionálisan az olvasó neve is feltüntethető, természetesen csak akkor, ha ez egyszerűen rendelkezésre áll, másrészt az olvasó ezt kifejezetten igényli. Amennyiben a dokumentumhoz való hozzáférés egyébként jogosítottsághoz kötött, valamint a letöltés DRM védeltséget kíván meg, úgy az egyedi azonosító „social DRM”-ként is szolgálhat.

Letöltés után az olvasó bármikor fordulhat a Loca Credibilia szolgáltatáshoz – www.locacredibilia.com

–, amelynek segítségével ellenőrizheti és bizonyíthatja a dokumentum hitelességét, azaz, hogy azt valóban az Országos Széchényi Könyvtár gyűjteményéből töltötte le és annak tartalma a letöltés óta nem változott. Ezt természetesen akkor is igazolja a Loca Credibilia rendszer, ha a szolgáltatási példány, amiről a felhasználói példány létrejött már nem érhető el a gyűjteményben, vagy időközben megváltozott volna. Amennyiben a Loca Credibilia rendszer az ellenőrzésre nemleges választ ad, úgy az Országos Széchényi Könyvtár mondhatja joggal, hogy a felhasználói példány nem az ő gyűjteményéből származik, annak tartalmaért nem tartozik felelősséggel.

A prototípus elkészült, annak integrálása az országos könyvtári platformmal folyamatban van. Sikeres integráció után az Országos Széchényi Könyvtár képes lesz bármilyen forrásból származó digitális dokumentumot – legyen az saját digitalizálás, más közgyűjteményben végzett digitalizálás, digitális kötelempéldány vagy egyéb forrás – bekerüléskor hitelesíteni, egyedileg példányosítani és hitelesen szolgáltatni. A Loca Credibilia szolgáltatás technikailag természetesen nemcsak az OSZK számára, hanem bármely digitális gyűjtemény számára elérhető.

Az új szolgáltatással reményeink szerint olyan eszközt adunk a könyvtárak, közgyűjtemények kezébe, amely megerősíti kivételes helyzetüket a digitális információforrások között, azt a hitelességet, amelyet megbízhatóságukkal vívtak ki, a technika segítségével ellenőrizhetővé is tehetik, sőt a tőlük származó digitális dokumentumpéldányok hitelességét a további felhasználás során (pl. tudományos hivatkozás vagy más bizonyítási eljárás) is igazolni tudják. Más bizalmi szolgáltatásoktól eltérően a Loca Credibilia nem csak azt képes igazolni, hogy egy dokumentum önmagában hiteles-e, de azt is, hogy melyik szolgáltatási példány másolatként jött létre, valamint akár azt is, hogy részei vagy egésze mely más példányokban kerültek felhasználására.

Irodalom

Az elektronikus források utolsó megtekintése: 2019. szeptember 23.

(EP3491575) BOSZNAY, Á. – ECKHARDT, P. – LANTOS, M. – SÁR, Cs. – SÜTHEŐ, P.: Method and system for the authentic determination of the identity of an electronic document with itself at a later date or with a copy thereof. Forrás: https://worldwide.espacenet.com/publicationDetails/biblio?DB=EPODOC&I=0&N=D=3&adjacent=true&locale=en_EP&FT=D&date=20190605&CC=EP&NR=3491575A1&KC=A1

- (ARL, 1995) Association of Research Libraries (ARL): Definition and purposes of a Digital Library. 1995. Forrás: <http://old.arl.org/resources/pubs/mmproceedings/126mmappen2~print.shtml>
- (Arms, 1997) ARMS, W.Y. – BLANCHI, C. – OVERLY, E. A.: An architecture for information in digital libraries. = D-Lib Magazine, February 1997.
Forrás: <http://mirror.dlib.org/dlib/february97/cnri/02arms1.html>
- (ARMS, 2000) ARMS, W. Y.: Digital library. Cambridge, MIT Press, 2000. 187 p.
- (Belik, 1999.) BELKIN, Nicholas: Understanding and supporting multiple information seeking behaviors in a single interface framework. Proceedings of the Eighth DELOS Workshop: User Interfaces in Digital Libraries. European Research Consortium for Informatics and Mathematics, 1999. 11–18. p.
- (Berners-Lee, 1992) BERNERS-LEE, Tim: World Wide Web. Forrás: <http://www.w3.org/pub/WWW/History/19921103-hypertext/hypertext/WWW/TheProject.html>
- (Buckland, 1997) BUCKLAND, Michael K.: What is a 'document'? = Journal of the American Society for Information Science, 48. vol. 9. no. 1997. 804–809. p.
- (Bush, 1945) Bush, Vanevar: As we may think. = The Atlantic Monthly, July, 1945.
- (DLI) National Science Foundation. Digital Libraries Initiative. Available Research, US Federal Government, 1999. Forrás: <http://dli2.nsf.gov/dlione/>
- (DLRM) CANDELA, L. – NARDI, A. (eds.): The Digital Library Reference Model (DLRM). DL.org. Coordination Action on Digital Library Interoperability, Best Practices & Modelling Foundations. Final Report 273 p. http://bscw.research-infrastructure.eu/pub/bscw.cgi/d222816/D3.2b_Digital_Library_Reference_Model.pdf
- (FRBR HUN): A bibliográfiai tételek funkcionális követelményei. (Functional Requirements for Bibliographic Records). Zárójelentés. (Final report). (FRBR HUN) Készítette az IFLA Bibliográfiai Tételek Funkcionális Követelményei. Munkacsoportja. (IFLA Study Group on the Functional Requirements for Bibliographic Records). Forrás: <https://www.ifla.org/files/assets/cataloguing/frbr/frbr-hu.pdf>
- (FRBR LRM) PAT RIVA, Pat – LE BOEUF, Patrick – ŽUMER, Maja: FRBR Library Reference Model. (FRBR LRM) Forrás: <https://www.ifla.org/publications/node/11412>
- (FRBR) Functional Requirements for Bibliographic Records. Final Report. (FRBR) UBCIM Publications. New Series. 19. vol. München, Saur, 1998. 136 p.
- (i2010) i2010: Digital libraries. Communication from the Commission of 30 September 2005 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – i2010: digital libraries [COM(2005) 465 final – Official Journal C 49 of 28.2.2008] Forrás: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:I24226i>
- (Kahn-Wilensky, 1995) KAHN, R. – WILENSKY, R.: A Framework for Distributed Digital Object Services. May 1995. Forrás: <http://WWW.CNRI.Reston.VA.US/home/cstr/arch/k-w.html>
- (Kunze, 2013) KUNZE, J. – RODGERS, R.: The ARK Identifier Scheme. Forrás: <https://www.ietf.org/archive/id/draft-kunze-ark-18.txt>
- (Leiner, 1998) LEINER, B.M.: The NCSTRL approach to open architecture for the confederated digital library. = D-Lib Magazine, 1998. Forrás: www.dlib.org/dlib/december98/leiner/12leiner.html
- (MacCall, 1999) MACCALL, S. L. – CLEVELAND, A. D. – GIBSON, I. E.: Outline and preliminary evaluation of the classical Digital Library Model. In: Knowledge, Creation, Organization and Use. Proceedings of the 62nd Annual Meeting of the American Society of Information Science, Washington, D.C. 1999. 33–46. p.
- (Nelson, 1965) NELSON, Ted: Complex information processing. A file structure for the complex, the changing and the indeterminate. In: Proceeding ACM ,65 Proceedings of the 1965 20th National Conference, Cleveland, Ohio, USA – August 24–26, 1965. ACM, New York, NY. 1965. 84–100. p.
- (Rakesh Kumar, 2016) RAKESH. Kumar Mishra: Digital libraries. Definitions, issues, and challenges. Innovare. = Journal of Education, 4. vol. 3. no. 2016. 1–3. p. Forrás: <https://innovareacademics.in/journals/index.php/ijoe/article/5665.html>
- (RFC 1630) Berners-Lee, T.: Universal Resource Identifiers in WWW. IETF RFC 1630. Forrás: <https://www.ietf.org/rfc/rfc1630.txt>
- (RFC 1738) BERNERS-LEE, Tim – MASINTER, L. – McCAHILL, M.: Uniform Resource Locators (URL). IETF RFC 1738. Forrás: <https://www.ietf.org/rfc/rfc1738.txt>
- (RFC 3986) BERNERS-LEE, T. – FIELDING, R. – MASINTER, L.: Uniform Resource Identifier (URI): Generic Syntax. IETF RFC 3986 Forrás: <https://www.ietf.org/rfc/rfc3986.txt>
- (RFC 8141) SAINT-ANDRE, P. – KLENSIN, J.: Uniform Resource Names (URNs). IETF RFC 8141. Forrás: <https://tools.ietf.org/html/rfc8141>
- (Seadle – Greifeneder, 2007.) SEADLE, Michael – GREIFENEDER, Elke: Defining a digital library = Library Hi Tech, 25. vol. 2. no. 2007. 169–173. p. Forrás: <http://www.emeraldinsight.com/doi/full/10.1108/07378830710754938>
- (Shiri, 2003) SHIRI, A.: Digital library research. Current developments and trends = Library Review, 52. vol. 5. no. 2003. 198–202. p.
- (Sütthő, 2001) SÜTTHŐ, Péter: Elektronikus, digitális, virtuális könyvtárak. In: Horváth Tibor, Papp István (szerk.) Könyvtárosok kézikönyve. 3. Budapest, Osiris, 2001. 209–240.p.
- (Yin-Leng Theng, 2002) YIN-LENG Theng: Information therapy in digital libraries. In: Digital libraries. People, Knowledge, and Technology. 5th International Conference on Asian Digital Libraries, ICADL 2002, Singapore, December 11–14, 2002; Berlin, Springer, 2002. 452–464. p.

Beérkezett: 2019. szeptember 26.

Melléklet: Definíciók a digitális könyvtár fogalmának meghatározásához

ARL, 1995:

„A digitális könyvtár nem egy egyedülálló entitás; A digitális könyvtár több erőforrás összekapcsolását lehetővé tevő technológiát igényel; A számos digitális könyvtár és információs szolgáltatás között a kapcsolatok transzparenssek a végfelhasználók számára; A digitális könyvtárakhoz és információs szolgáltatásokhoz való egyetemes hozzáférés a célkitűzés; A digitális könyvtár nem korlátozódik a dokumentum-szurrogátumokra, hanem a nyomtatásban nem megtestesíthető és terjeszthető műtárgyakra is kiterjed.”

Arms, 2000:

„Nem hivatalos definíció szerint a digitális könyvtár információk rendszerezett és kezelt gyűjteménye, melyhez szolgáltatások kapcsolódnak, és ahol az információt digitális formában tárolják és hálózatok útján teszik hozzáférhetővé. E meghatározásban kiemelt szerepe van az információ kezelésének. A műholdról a földre küldött adatfolyam nem könyvtár. Ugyanazok az adatok, ha szisztematikusan rendszerezve vannak, digitális könyvtári gyűjteménnyé válnak. A legtöbb ember nem tekintene digitális könyvtárnak egy adatbázist, mely egyetlen vállalat pénzügyi adatait tartalmazza, de több vállalat ilyen típusú információinak gyűjteményét már el tudja fogadni egy könyvtári gyűjtemény részeként. A digitális könyvtárak különféle információkat tartalmaznak, sokféle felhasználó számára. A digitális könyvtárak mérete az egészen kicsitől a hatalmas méretűig terjed, és bármilyen típusú számítógépes berendezést és bármilyen megfelelő szoftvert használhatnak. Ami egységessé teszi őket az az, hogy az információkat számítógépen szervezik és hálózaton keresztül teszik elérhetővé a következő eljárások segítségével: az anyag kiválasztása a gyűjteményben; annak rendezése; azok elérhetővé tétele a felhasználók számára; valamint azok archiválása.”

MacCall, 1999:

„Egy digitális könyvtár olyan elektronikus tudásforrás gyűjtemények gyűjteménye, amelyet annak érdekében fejlesztettek ki és tartanak fenn, hogy elegendő legyen egy adott felhasználói kör minden információigényének.”

Wikipedia (az angol nyelvű oldal szövegének fordítása):

„A digitális könyvtár egy speciális, digitális objektumok gyűjteményét tartalmazó könyvtár, mely magába foglal szöveget, vizuális anyagot, audio- és videófelvételt elektronikus médiaformátumban tárolva (szemben a nyomtatott, vagy más médiaformátumokkal), valamint eszközöket azok szervezésére, tárolására és letöltésére. A digitális könyvtárak méretüket és terjedelmüket tekintve nagyon különbözhetnek, a fenntartójuk lehet magánszemély, vagy szervezet, esetleg kapcsolódhatnak létező könyvtári létesítményekhez, vagy szervezetekhez, vagy tudományos intézményekhez ill. intézetekhez. A digitális tartalmat tárolhatják helyben, vagy távolról elérhető számítógépes hálózatokon. Az elektronikus könyvtár egyfajta információkereső rendszer. Ezek az információkereső rendszerek képesek egymással információt cserélni az átjárhatóság és fenntarthatóság érdekében.”

DLI:

„A digitális könyvtárak alapvetően elektronikus dokumentumokat tárolnak, és az ezekből képzett nagy gyűjteményeket hatékonyan

kezelik. A digitális könyvtárak felkeresése olyan hálózati információs rendszerek felkeresése, amelyek arra összpontosítanak, hogy miként lehet fejleszteni a szükséges infrastruktúrát az interneten található információk hatékony és tömeges kezeléséhez.”

Seadle – Greifeneder, 2007:

„A digitális könyvtár alapvetően olyan erőforrás, amely digitális formában rekonstruálja a hagyományos könyvtár szellemi tartalmát és szolgáltatásait. A digitális könyvtárak magukban foglalják a digitális tartalmakat (amelyek néha, de nem feltétlenül szövegalapúak), kapcsolatokat (amelyek lehetnek egyszerű linkek vagy összetett metaadatok vagy lekérdezés-alapú kapcsolatok) és a szoftvereket (amelyek lehetnek egyszerű HTML oldalak vagy összetett adatbázis-kezelő rendszerek). Egyetlen, egyszerű, önmagában álló weboldal valószínűleg semmilyen értelemben nem digitális könyvtár, mint ahogy egyetlen oldal, vagy egy könyv sem alkot önmagában hagyományos könyvtárat. A nyers adatok halmaza, amelynek például a Hubble űrtéleszköpről érkeznek, valószínűleg szintén nem alkotnak digitális könyvtárat, noha tartalmuk miatt vitathatatlanul összetartoznak. A digitális könyvtárak nem helyettesítik a hagyományos könyvtárakat. Inkább a tradicionális könyvtárak jövőjét jelentik, ugyanúgy, ahogy a középkori kéziratok könyvtárak egyszerűen a mai nyomtatott könyvtárak speciális és megbecsült részévé váltak.”

Leiner, 1998:

„A digitális könyvtár szolgáltatások és információs objektumok gyűjteménye, mely támogatja a felhasználókat az elektronikus/digitális eszközökkel közvetlenül vagy közvetetten elérhető információs objektumok kezelésében.”

Shiri 2003:

„Azok a szervezetek, melyek forrásokat biztosítanak, ideértve a speciális személyzetet is a digitális művek gyűjteményének kiválasztásához, felépítéséhez, a hozzáférés biztosításához, értelmezéséhez, terjesztéséhez, az integritás megőrzéséhez, valamint a folyamatos elérhetőség biztosításához, annak érdekében, hogy ezek a források könnyen elérhetőek legyenek egy meghatározott közösség vagy több közösség számára.”

Rakesh Kumar 2016:

„A digitális könyvtár egy szolgáltatás-gyűjtemény, információs objektumok gyűjteménye, amely a felhasználókat látja el információk objektumokkal, ezen objektumokat szervezi és megőrzi, közvetlen vagy közvetett elérhetőséggel, és elektronikus/digitális hozzáféréssel. A digitális könyvtár elsődleges célja a hozzáférés javítása, csakúgy mint a költség-hatékony megőrzés, a lépés tartása a technikai fejlődéssel, valamint az információ megosztás.”

i2010:

„A digitális könyvtárak a nyilvánosság számára elérhető digitális tartalom rendszerezett gyűjteményei. A tartalom olyan anyag, amit vagy digitalizáltak (könyvek és egyéb dokumentumok másolatai), vagy amelyet eredetileg digitális formátumban készítettek.

Három kiemelt terület van, ahol a digitális technológiák lehetőségeit ki kell aknázni az információkhoz való hozzáférés bővítése érdekében:

- online hozzáférhetőség;
- az analóg gyűjtemények digitalizálása;
- a digitális tartalom megőrzése és tárolása.”