

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

Bölcsészettudományi Kar

Szakirányú továbbképzés szakdolgozat

HOSSZÚ TÁVÚ DIGITÁLIS MEGŐRZÉS A HAGYOMÁNYOK HÁZA MÉDIAARCHÍVUMÁBAN

Tervek a Folklór Archívum
gyűjteménykezelési folyamatának
szabványokon alapuló megújítására

Témavezető:

Alföldi István

Palkó Gábor, tanszékvezető

Készítette:

Zagyva Natália

Adatgazdász szakirányú továbbképzés

Tartalomjegyzék

1.	Bevezetés	3
2.	A Hagyományok Háza Folklór Archívuma	5
2.1	Intézményi háttér.....	5
2.2	Gyűjtőkör	7
2.3	A gyűjtemény fizikai és digitális objektumai.....	8
2.4	A gyűjteménykezelés folyamata	9
2.5	Megoldandó problémák	10
3.	Archívumi szabványok	14
3.1	OAIS	14
3.2	PREMIS	18
3.3	METS	20
3.4	EAD.....	21
3.5	ISAD(G).....	21
3.6	Dublin Core	22
3.7	OAI-PMH.....	23
4.	Nemzetközi kutatások és tapasztalatok.....	25
4.1	E-ARK program	25
4.2	ISA ² program	29
5.	Nyílt forráskódú és szabványos megoldásokat kínáló szoftverek	31
5.1	A nyílt forráskódú szoftverek előnyei és hátrányai.....	31
5.2	Gyűjteménykezelő rendszerek	31
5.3	Hosszú távú megőrzésre szolgáló rendszerek	34

6.	Eredmények és tervek	36
7.	Irodalomjegyzék	37
8.	Mellékletek	40
8.1	Összegzés	40
8.2	Summary	41

1. BEVEZETÉS

A Hagyományok Házában 2023 elején nagy szervezeti átalakítás történt, melynek eredményeképp létrejött egy új szervezeti egység Adatbázis Szakcsoport néven. A szakcsoport néhány – témánk szempontjából releváns – feladata a Hagyományok Háza Szervezeti és Működési Szabályzata [1] szerint:

- az intézmény adatbázisainak fejlesztése, egységes rendszerbe szervezése, összehangolása;
- a Hagyományok Háza különböző részlegeinek működése során használt, már meghaladott adatbázis-technológiák kivezetése, egyidejűleg a feladatoknak leginkább megfelelő korszerű digitalizációs megoldások alkalmazásának elősegítése;
- szükséges új adatbázisokra vonatkozó igények és szakmai indokok felmérése;
- az adatbázisokkal kapcsolatos informatikai fejlesztések tervezése, specifikáció elkészítése, a fejlesztés megvalósításának koordinálása, tesztelés, kapcsolattartás a külső fejlesztőkkel;
- a munkatársak támogatása a megvalósult adatbázisok használatában.

2024-ben a szakcsoport egyik legfontosabb projektje az intézmény digitális médiaarchívumának megtervezése a Folklór Archívum gyűjteményének vonatkozásában, mely tartalmaz analóg, digitalizált és digitálisan keletkezett (*born digital*) anyagokat is. A cél, hogy a jelenleg polcokon, hálózati mappákban és külső merevlemezekben tárolt, valamint a későbbiekben beérkező vagy elkészülő archiválendő képeket, videó- és hangfelvételeket, kéziratokat, valamint a hozzájuk tartozó dokumentációt egységes rendszerben tudjuk kezelni, metaadatokkal ellátni, kereshetővé és elérhetővé tenni, és a hosszú távú megőrzésük is biztosítva legyen. A tervezett médiaarchívum egyaránt szolgálja majd a múlt és a jelen kutatását, a folklór- és a folklórizmuskutatást. Terveink szerint a médiaarchívumon alapul majd a folklórgyűjtések részletes tartalmi szegmentálását és adatolását tartalmazó Folklóradatbázis [2] új szerkesztőrendszere, valamint a készülő, jelenleg adatgyűjtés fázisában lévő szakmatörténeti archívum. A Hagyományok Háza operatív működését is hatékonyabbá

teheti azzal, hogy a marketingfolyamatokhoz könnyen megtalálható és elérhető médiaanyagokat biztosít.

A hosszú távú megőrzés módszerei ma már szinte minden esetben az Open Archival Information System (OAIS) referenciamodellre épülnek, ezért munkánk során mi is erre a modellre épülő megoldást keresünk. Lényeges szempont az is, hogy az új rendszer jól és könnyen integrálható legyen más szoftverekkel, és természetesen a gazdaságos, hosszú távú fenntarthatóság is alapvető elvárás. A szakdolgozat összefoglalja a digitális archiválást segítő sztenderdeket, módszereket és jó gyakorlatokat, és az ezeket megvalósító informatikai megoldások összehasonlításával megkísérli letenni a tervezett médiaarchívum szoftverarchitektúra-tervének alapjait.

2. A HAGYOMÁNYOK HÁZA FOLKLÓR ARCHÍVUMA

2.1 Intézményi háttér

A Hagyományok Háza 2001-ben jött létre, de a jelenleg Folklór Archívumnak nevezett audiovizuális archívum története ennél korábbra nyúlik vissza. Az előzményekben megidézhetnénk a magyar népzene, néptánc, bővebben a folklór és folklorizmus kutatóit, intézményeit, de ebben a fejezetben csak a Hagyományok Háza székházához, a Budai Vigadó épületéhez kötődő intézményi előzményeket, népművészeti és közművelődési tevékenységeket foglalom röviden össze. Bár nem maga az épület a fontos, az élet mégis úgy hozta, hogy az itt történtek jelentősen befolyásolták a Folklór Archívum létrejöttét és gyűjtőkörének kialakulását.

Az 1951-ben a szovjet „Népi Alkotások Háza” mintájára alapított Népművészeti Intézetnek a Budai Vigadó adott otthont. Az új intézmény vezetője, Széll Jenő sikeresen bevonta a munkába azokat a szakembereket, akik a néhány évvel korábban felszámolt amatőr művészeti-művelődési szövetségekben dolgoztak. A népzene- és néptánckutatás fontos szereplői – többek közt Muharay Elemér, Martin György, Pesovár Ernő, Pesovár Ferenc – lehetőséget kaptak az intézetben, hogy a közművelődési tevékenység mellett elméleti, módszertani, tudományos kérdésekkel is foglalkozzanak [3]. A néptáncsal és népzenevel foglalkozó szakemberek fontosnak tartották, hogy gyűjtéseket végezzenek és ezeket minél szélesebb körben – az amatőr néptáncmozgalom számára is – hozzáférhetővé tegyék. Domokos Péter Pál 1952-ben született beszámolója szerint már a megalakulás évében 30 gyűjtő, 30 gyűjtőúton csaknem 4000 népdalt gyűjtött. A gyorsan bővülő gyűjtemény rendezésére és archiválására is nagy gondot fordítottak [4]. Az 1956-os teljes átszervezés után 1957-től Népművelési Intézet néven működő intézményben tovább folytatódott a néprajzi gyűjtőmunka, megerősödött a néptáncosok tanfolyami és az amatőr együttesek minősítési rendszere. Ebben a korszakban is sok olyan táncos, zenész, koreográfus dolgozott itt, akiknek tevékenysége meghatározó a Folklór Archívum későbbi gyűjteményére nézve, például Manga János, Szigeti Károly, Györgyfalvy Katalin, Novák Ferenc, Kricskovics Antal, Timár Sándor, Vásárhelyi László. A hetvenes években a Vitányi Iván által vezetett intézmény felkarolta és szakmailag is

segítette az akkor kibontakozó táncházmozgalmat [5]. 1982-ben az ekkor már Művelődési Intézetnek nevezett intézmény keretében létrejött a Néptáncosok Szakmai Háza, mely részt vett a tanfolyamok és fesztiválok szervezésében, foglalkozott oktatási segédanyagok kiadásával, gyűjtötte és szolgáltatta az érdeklődőknek az archívumában – gyakran másolatban – őrzött gyűjtések és az amatőr néptáncgyűttesek által előadott műsorok, fesztiválok felvételeit [6].

A Magyar Állami Népi Együttes története is 1951-ben kezdődött. Az ekkor alapított, Rábai Miklós által vezetett hivatásos együttes tánckarból, énekkarból és cigányzenekarból állt, népzenei és néptánc ihletésű színpadi előadásokat mutatott be többféle műfajban. A hetvenes években a táncházmozgalom hatott a színpadi néptáncra is. Timár Sándor vezetésével a Magyar Állami Népi Együttes is középpontba állította az eredeti népzene és néptánc bemutatását. Az ehhez szükséges elmélyült és alapos tudást csak archív felvételek és további gyűjtések alapján lehetett megszerezni.

Sebő Ferenc, aki korábban a Népművelési Intézet kutatási osztályán, majd a Zene-tudományi Intézetben is dolgozott, a kilencvenes évek második felében a Magyar Állami Népi Együttes művészeti vezetőjeként fontosnak tartotta, hogy létrejöjjön egy olyan tudományos igényességű folklórarchívum a hivatásos együttes keretein belül, amely az amatőr táncosok és zenészek számára is szolgálatot. Felkérte Pávai István népzene-kutatót, hogy szervezze meg és vezesse ezt a műhelyt. Így 1999-ben új munkatársakkal és eszközparkkal, valamint Martin György közel hatezer kötetes könyvtárával, kéziratos hagyatékaival és hangfelvételeivel létrejött a Hagyomány-őrző Műhely és a Martin Médiatár a Budai Vigadó épületében. Az archívum szükségességének indokairól így írt akkoriban Pávai a Folkmagazin hasábjain: *„A hetvenes évek elején beinduló, majd napjainkig egyre terebélyesedő táncházmozgalom az eredeti folklóranyag iránti érdeklődést a korábbi időszakhoz képest sokszorosára emelte, ami egybevágott a viszonylag olcsó és könnyen hozzáférhető kazettás hang-, majd mozgóképrögzítő berendezések elterjedésével. A népzenevel és néptáncsal kapcsolatos táncvárosi és színpadi munka szempontjából ezek az eszközök előnyösnek bizonyultak, bárki létrehozhatott a lakásán kisebb-nagyobb folklór-archívumot.”*

chívumot. A tudományos feldolgozás, a hosszú távú biztonságos megőrzés, a közérdekű reprodukálhatóság tekintetében viszont nem kínáltak megnyugtató megoldásokat” [7].

Két évvel később, 2001-ben – részben a fent említett intézmények összevonásával – létrejött a Hagyományok Háza elnevezésű új nemzeti intézmény három nagy egységgel, a művészeti tevékenységet folytató Magyar Állami Népi Együttessel, a közművelődési tevékenységet folytató Népművészeti Módszertani Műhellyel és a gyűjteményi tevékenységet folytató Lajtha László Folklórdokumentációs Központtal. Az azóta több szervezeti átalakulást és névváltozást megélt folklórdokumentációs archívum jelenleg Folklór Archívum néven a Hagyományok Háza Gyűjteményi Főosztálya alá tartozik.

2.2 Gyűjtőkör

Az archívum gyűjtőkörét elsősorban az előző fejezetben leírt előzmények és igények határozták meg. Földrajzi tekintetben a fő gyűjtőkör kezdetben és most is a magyar nyelvterület. A Kárpát-medencén és Moldván kívüli területek közül a szomszéd és rokon népek hagyományos kultúrája tartozik még az archívum látóterébe. Eredetileg a központi tevékenység a folklórgyűjtések digitalizálása, archiválása és feldolgozása volt. Jelenleg is ez áll első helyen, de egyre több új, intézmény- és szakmatörténeti, a folklorizmus jelenségeit dokumentáló anyag kerül az archívum polcaira, a feldolgozásuk pedig a teendők listájára. A Folklór Archívum tárolja és tartja nyilván a Halmos Béla által megálmodott és létrehozott Táncház Archívum dokumentumait (újságcikkek, kéziratok, aprónyomtatványok, fotók, hang- és videófelvételek, interjúk, oktatási anyagok stb.), a MÁNE archívumát (kották, plakátok, fotók, hang- és videófelvételek stb.) és több hagyatékot (Halmos Béla, Lajtha László, Martin György, Kallós Zoltán, Erdélyi Zsuzsanna, Farkas Gyula). A 2021-ben kiadott Gyűjtőköri és Állományalakítási Szabályzat a fentieknél részletesebben és bővebben, a Martin György Szakkönyvtárra is kiterjesztve tárgyalja a gyűjtőköri szinteket és az állománygyűjtés szempontjait [8]. Óriási kihívást jelent az archívumnak az az elvárás, hogy a jelen eseményeinek dokumentumait is megőrizze. Ez az igény két irányból érkezik: egyrészt a folklorizmus jelenségeit, a táncházmozgalom történetét rögzíteni, kutatni és megismerni vágyó, másrészt a marketing tevékenységet végző, illusztrációs anyagokat kereső munkatársak felől. Jól átgondoltan,

körültekintően kell a kapacitásokkal gazdálkodni, a prioritásokat meghatározni ahhoz, hogy a jelenben keletkező rengeteg digitális dokumentum és a múltban született archív anyagok feldolgozásának aránya ne boruljon fel.

A Folklór Archívumban a gyűjteményi egységek rendszerét többnyire a forrásarchívumok határozzák meg. A forrásarchívumok nagy része magángyűjtemény, de vannak szép számban intézmények is a beadók közt, például a Csíkszeredai Regionális Archívum, az Erdélyi Néprajzi Múzeum, a Folk Center Alapítvány, a Folkrádió, a Fonó Budai Zeneház, a Fővárosi Szabó Ervin Könyvtár, a Hagyományok Háza előzményeihez kötődő intézmények (MÁNE, Néptáncosok Szakmai Háza), a Hargita Megyei Kulturális Központ, a Háromszék Táncegyüttes, a Jászság Népi Együttes, a Kanizsai Dorottya Múzeum, a Mária Rádió (Erdély), a Maros Művészegegyüttes, a Muharay Elemér Népművészeti Szövetség, a Népművészeti Egyesületek Szövetsége, a Néprajzi Kutatóintézet, a Néprajzi Múzeum, a Nyírség Táncegyüttes, a Szentendrei Szabadtéri Néprajzi Múzeum, a Székelykeresztúri Múzeum és a Zene tudományi Intézet. Említsünk meg néhány ismertebb nevet a legtöbb anyagot beadott magángyűjtők közül: Agócs Gergely, Balázs Gusztáv, Barabás László, Bodó-Bán János, Dénes Zoltán, Fazakas János, Halmos Béla, Havasréti Pál, Juhász Zoltán, Kelemen László, Kerényi Róbert, Könczei Árpád, Nesztor Iván, Pávai István, Sára Ferenc, Sáringer Kálmán, Sebő Ferenc, Stuber György, Timár Sándor, Varga Sándor, Vavrinecz András, Vucic Gyula.

2.3 A gyűjtemény fizikai és digitális objektumai

A Folklór Archívum sok olyan audiovizuális hordozót őriz, amely műtárgyként nem értékes. Az érték nem maga a fizikai objektum, hanem a rajta tárolt információ, ami jó minőségű digitalizálás után és biztonságos, hosszú távú megőrzést biztosítva teljes értékű az adathordozó fizikai létezése nélkül is.

Az archívum – a biztonságos hosszú távú megőrzés jelenleg meglévő hiányosságai miatt is – törekszik arra, hogy ha a gyűjteménytulajdonos nem kéri vissza, akkor tárolja a beérkezett hordozókat is. Az elmúlt 25 évben már többször előfordult, hogy a digitális fájl sérülése miatt vagy jobb minőségű, esetleg más formátumú eredmény érdekében szükség volt bizonyos anyagok újradigitalizálására. Néhány

éve sikerült beszerezni egy korszerű filmdigitalizáló eszközt, így a 10-20 éve gyengébb technikával digitalizált, de értékes tartalmú filmeket érdemes lenne újradigitalizálni.

A gyűjteményben egyre több a *born digital* média, melyek gyakran nem az eredeti hordozón, hanem digitális másolati példányként kerülnek be az archívumba. A Folklór Archívum történetében még nem volt arra példa, hogy olyan digitális hagyatékkal kellett volna dolgozni, ahol a teljes informatikai környezetet, például egy laptopon tárolt összes fájlt, levelezést változatlan állapotban menteni kellett volna.

Az archívum klimatizált raktárában tárolt hordozók típusai: fényképek, transzparens hordozók (dia, negatív, film), mágneses hordozók (orsós és kazettás magnó- és videoszalagok), hanglemezek, optikai hordozók, mágneses lemezek (HDD), szilárdtest-meghajtók (pendrive, memóriakártya, SSD), valamint az ezekkel tematikusan szorosan összefüggő kéziratok [9].

2.4 A gyűjteménykezelés folyamata

A Folklór Archívum gyűjteménykezelési folyamatának lépései a Múzeumi Spektrum [10] fogalomrendszerével leírva röviden, pontokba szedve:

Gyűjteménygyarapítás

- Előzetes feltárás
- Értékbecslés
- Szerződés előkészítése
- Szerződéskötés
- Átadás-átvétel
- Állapotfelmérés, szükség esetén műtárgyvédelmi tevékenység

Nyilvántartás, feldolgozás

- nyilvántartásba vétel a feltárás aktuális szintjén
- fizikai objektum (kézzel fogható tárgy) esetén:
 - elhelyezés a gyűjteményi raktárban
 - digitalizálás

- digitális objektumok kezelése (archiválendő példányok kiírása külső merevlemezre, felhasználói igényektől függő használati reprezentációk készítése és elérhetővé tétele)
- katalogizálás (adatolás, adatellenőrzés, adatjavítás, tezaurusz karbantartása)

Megőrzés

- fizikai objektum esetén:
 - állományvédelem
 - biztonságos raktározás
- a digitális objektumok hosszú távú megőrzését elősegítő tevékenységek

Hozzáférés biztosítása

- valós térben (kiállítás, kutatószolgálat)
- digitális térben (online adatbázisok, digitális tartalmak, virtuális kiállítások)

2.5 Megoldandó problémák

Ebben a fejezetben azokat a legfájóbb pontokat sorolom fel, melyek miatt az archívum- és az intézményvezetés prioritásként kezeli az új folyamatok és eszközök bevezetését. A megoldások keresésekor figyelembe kell venni az emberi és anyagi erőforrások szükségességét és bővítési lehetőségeinek korlátait.

Gyűjteménygyarapítás

A *born digital* anyagok beadásának még nincs kiforrott, szabályozott, automatizált folyamata és felülete. Elsőként az intézmény munkatársai számára kellene biztosítani egy olyan lehetőséget, ahol rögtön a létrehozásuk és kiválogatásuk után, leíró adatokkal együtt könnyedén benyújthatják az általuk rögzített digitális anyagokat, például a Hagyományok Háza rendezvényein készített felvételeket.

Nyilvántartás, feldolgozás

Jelenleg a nyilvántartás Excel-táblákban történik. Minden forrásgyűjteményhez készül egy táblázat, általában annyi lappal, ahány gyűjteményi egység van benne. Ebben tartják nyilván a gyűjteménykezelők a gyarapodási, digitalizálási munkafázisok adatait és eredményét, de később az anyag részletes feldolgozásával foglalkozó katalogizálók is gazdagítják a tartalmat az általuk megállapított – és párhuzamosan a

Folklóradatbázisban is rögzített – leíró adatokkal. Ennek a módszernek rendkívül sok hátránya van. Néhány közülük: korlátozott hozzáférés, nehézkes kereshetőség, redundáns adattárolás, áttekinthetetlenség, sok hiány és hibalehetőség, nem egységes adatszerkezet. Egyértelműen az a legsürgősebben megoldandó feladata az archívumnak, hogy egy adatbázis alapú gyűjteménykezelő rendszert bevezessen, és a több száz Excel-táblában tárolt nyilvántartást minél teljesebb mértékben importálja. Az importálás a nem egységes adatszerkezetek és a redundáns adatok miatt bonyolult és időigényes munka lesz. Elképzelhető, hogy érdemes először egy kevés adatot rögzítő, de a teljes gyűjteményt átfogó importot készíteni, és később alapos adatellenőrzéseket követően, gyűjteményi egységenként bővíteni és gazdagítani a metaadatokat.

A Hagyományok Háza szeretné a folklórányagok részletes katalogizálására szolgáló, Folklór Archiváló Rendszer elnevezésű, eddig nagy szolgálatot tett, de mára több szempontból elavult, nehezen integrálható, több mint 20 éves egyedi asztali szoftvert alapjaiban megújítani. Az új informatikai rendszer tervezése során született meg az a döntés, hogy érdemes először a gyűjteménykezelést megújítani úgy, hogy a szoftverek jól tudjanak egymással kommunikálni, adatokat cserélni. A tervezett új gyűjteménykezelő és megőrzési rendszer belső használatú munkaneve Médiaarchívum lett. (A szakdolgozat címében ezt használom, de a jelenlegi tervek szerint nem ez lesz az új rendszer megnevezése.)

Megőrzés

Az archívum megalakulásakor és utána még hosszú évekig CD-re és DVD-re kerültek a megőrzendő nagy felbontású médiafájlok, mert nem volt másra lehetőség. Később már elérhető árúak lettek a merevlemezek, és elkezdtek átmásolni rájuk a CD-kre és DVD-kre mentett anyagokat. Sajnos volt köztük olyan, ami már nem volt olvasható, ezek egy nagy részét sikerült másolati példányokról pótolni. Jelenleg az archív fájlok egy példánya HDD-re kerül, másik példánya LTO-ra, de mindkettőt ugyanott, az archívum raktárában tárolják. Cél a georedundáns tárolás megoldása. A nem hálózatba kötött tárolóeszközök esetén megvan a kockázata annak, hogy nem veszi észre senki, amikor egy hordozó megsérül vagy tönkremegy, és a rajta tárolt anyag részben vagy egészben elérhetetlenné válik. A médiaarchívum

tervezési feladatai közé tartozik a helyi vagy távoli hálózati megoldások költségeinek felmérése. A merevlemezeken tárolt „mély archívum” mérete jelenleg körülbelül 320 TB. A folyamatos elérésű, hálózatra kötve tárolt, kis felbontású használati fájlok nagyságrendileg 10 TB helyet foglalnak.



1. ábra
Merevlemezek (felül) és LTO-k (alul a bőröndökben) a Folklór Archívum raktárában
Fotó: Káplár Béla

Nagy hiányossága még a jelenlegi rendszernek a szolgáltatásba kerülő digitális reprezentációk nem automatizált, verziókezelés nélküli létrehozása.

Hozzáférés biztosítása

Az archívum felfedezésére, kutatására, az őrzött digitális tartalmak elérésére jelenleg a következő lehetőségek vannak:

- A Folklór Archiváló Rendszerben részletesen katalogizált és publikusnak nyilvánított anyagok elérhetők bárki számára a Folklóradatbázis weblapján.
- A Folklór Archiváló Rendszerben részletesen katalogizált, de nem publikus anyagok kutatási engedéllyel elérhetők a Martin György Szakkönyvtárban.
- A már digitalizált, de nem feldolgozott anyagok egyedi elbírálással szolgáltatottak a Martin György Szakkönyvtárban.
- Válogatott, tematikus anyagok megjelennek kiállításokon és kiadványokban.

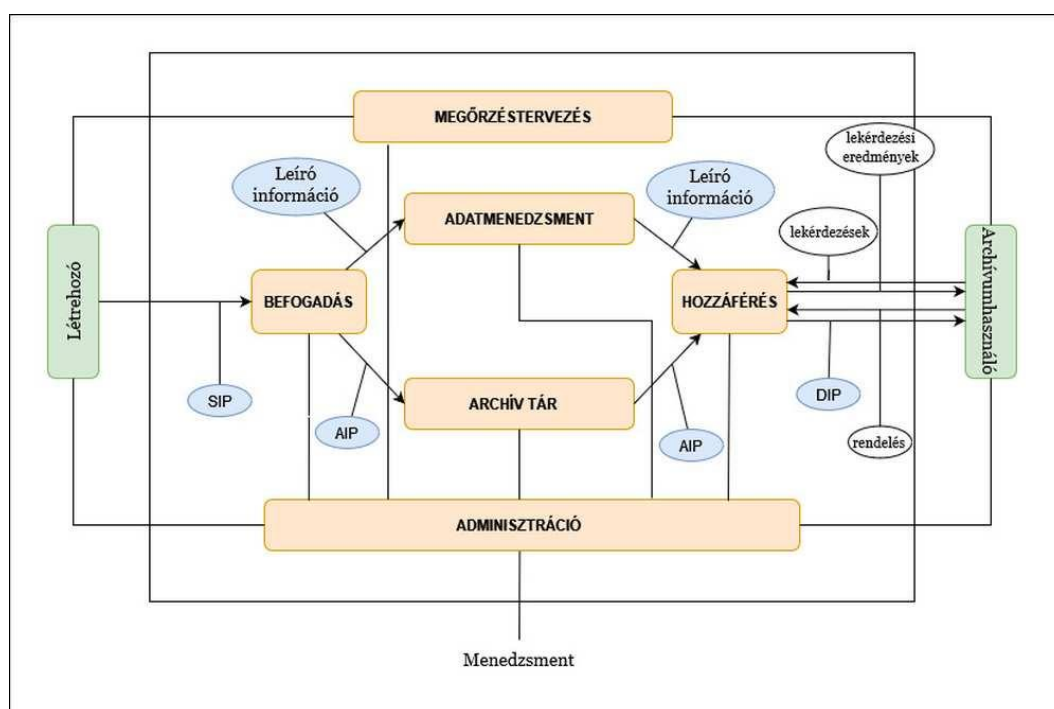
Az Excel-táblákban történő és nem szabványokon alapuló nyilvántartás nem teszi lehetővé, hogy szoftverek integrációja és szabványos adatcsere útján gépek is olvasni tudják az archívumban tárolt objektumok metaadatait, és ez nagyban akadályozza az emberek hozzáférését is. A kívánatos aggregációs szolgáltatások használata sem jöhet így szóba.

3. ARCHÍVUMI SZABVÁNYOK

Az archívumokra és közgyűjteményekre vonatkoztatható nemzetközi és nemzeti szabványok köre nagyon széles, közülük a dolgozatban csak azokat emelem ki, melyek a Folklór Archívum problémáira megoldást nyújtó szoftverek kiválasztása szempontjából a leglényegesebbek.

3.1 OAIS

Az Open Archival Information System (OAIS), magyarul Nyílt archívumi információs rendszer referenciamodell egy olyan átfogó, elméleti keretrendszer, amely segíti az archívumokat, hogy hosszú távon megőrizzék és hozzáférhetővé tegyék a digitális információkat. Az OAIS referenciamodell első változatát a NASA kezdeményezésére, eredetileg űrkutatási adatok megőrzésére dolgozták ki az 1990-es években. Először 2002-ben szabványosították, majd 2012-ben lépett hatályba a második, átdolgozott változata (ISO 14721:2012) [11].



2. ábra

Ilácsa Szabina és Bódog András (KöSZI) ábrája
az MSZ ISO 14721:2022 szabvány alapján [12]

Mivel a referenciamodell nem kötődik konkrét technológiához, hanem elvi, fogalmi kereteket fektet le, ezért általánosan használható az archívumok tervezésére, vizsgálatára és összehasonlítására. Mára széles körben ad iránymutatást a memóriaintézmények digitális archívumainak kialakításához is. Az OAIS referenciamodell nagy mértékben elősegítette a szakma közös szókincsének kialakítását, az archiválási fogalmak és funkciók egyértelműsítését és megértését.

Az OAIS szabványt 2022-ben Magyarországon is honosították [12] [13]. Az 2. ábra a magyar szabványban szereplő terminológiával mutatja be az OAIS sémát.

A modell meghatároz kulcsszereplőket, információs és funkcionális entitásokat.

Környezeti modell

Az OAIS referenciamodell azonosítja az archívum környezetét alkotó entitásokat:

- **Létrehozó (Producer):** A megőrzendő információt szolgáltató személy, szervezet vagy szoftver.
- **Archívum (Archive):** Az a szervezet, amely megőrzi és hozzáférhetővé teszi az információt a célközönség számára.
- **Archívumhasználó (Consumer) és célközönség (Designated Community):** Az archívumhasználók azok a személyek, szervezetek vagy szoftverek, akik kapcsolatba lépnek az archívummal a megőrzött információ elérése érdekében. A célközönség az előre megtervezett és megcélzott archívumhasználók halmaza, az a kijelölt közösség, amelynek számára az archívum az információt őrzi, érthetővé és értelmezhetővé teszi.
- **Menedzsment (Management):** Az a szervezet, amely meghatározza az archívumi stratégiát, gyűjtőkört, irányelveket, gyakran az archívum finanszírozásának forrása is egyben. Nem felel az archívum napi működésének irányításáért.

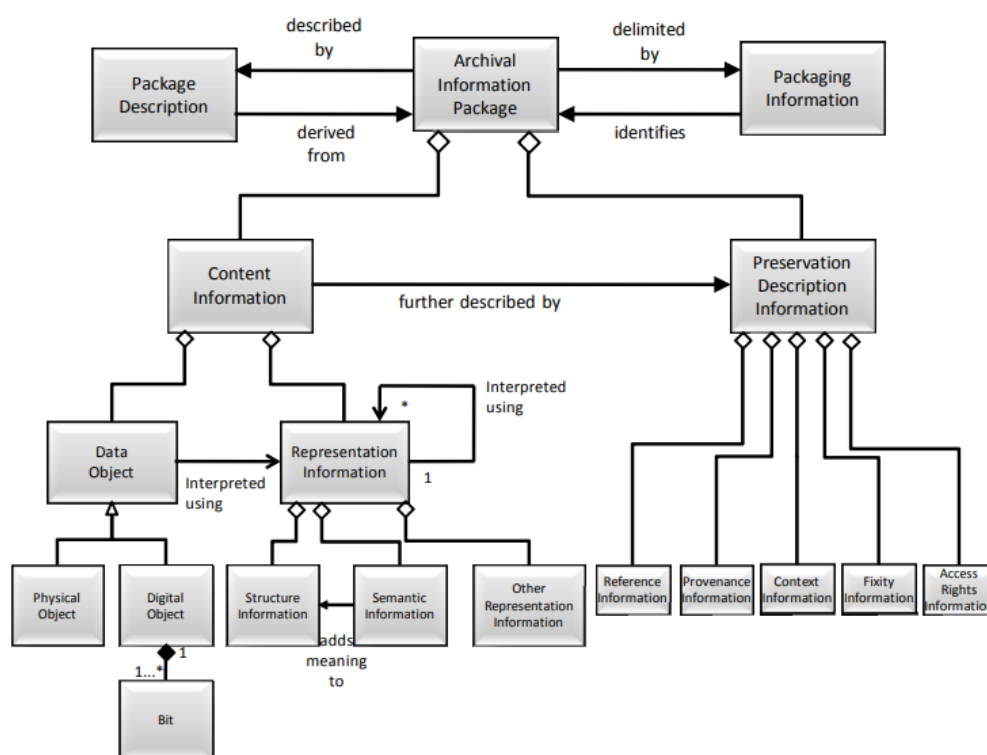
Információs modell

Az OAIS referenciamodell három különböző típusú, több fájlból álló információs csomagot definiál:

- **SIP (Submission Information Package):** A létrehozó által az archívumnak átadott adatcsomag

- AIP (Archival Information Package): Az archívumban tárolt és őrzött adatcsomag
- DIP (Dissemination Information Package): az archívumfelhasználó igényeinek megfelelően összeállított és szolgáltatott adatcsomag

Az archívum fő célja az AIP hosszú távú, a célközönség számára értelmezhető megőrzése, ezért erről a csomagról írok egy kicsit bővebben. Minden információs csomagnak tartalmaznia kell az adatot és a hozzá kapcsolódó leíró metaadatokat, de az AIP emellett még a tartalom hosszú távú megőrzését és az értelmezhetőséget lehetővé tevő metaadatokat is magában foglalja.



3. ábra
AIP [11]

A 3. ábra látható az AIP koncepcionális modelljének részletes nézete. A modell fő elemei:

- Tartalominformáció (Content Information)
 - Adatobjektum (Data Object): a megőrzésre szánt fizikai vagy digitális objektum. A digitális objektum állhat egy vagy több fájlból,

formátuma sokféle lehet, például szöveg, kép, hang, videó, adatbázis, szoftver vagy ezek kombinációja.

- Reprezentációs információ (Representation Information): segít megjeleníthetővé és értelmezhetővé tenni az adatobjektumot. Tartalmazhat például leírást arról a szoftverkörnyezetről, amely a megjelenítéshez szükséges vagy adhat egy összefoglalást az adat megfelelő értelmezéséről, az adatstruktúráról.
- Megőrzési leíró információ (Preservation Description Information)
 - Azonosító (Reference): a tartalominformáció egyedi azonosítója, lehet a rendszer által generált belső azonosító vagy egy külső egyedi azonosító (pl. ISBN, DOI)
 - Kontextus (Context): a tartalominformáció kapcsolatai más tartalominformáció-objektumokkal
 - Proveniencia (Provenance): dokumentálja a tartalominformáció történetét, minden tartalmi vagy formai változtatás időpontját, tényét, eredményét
 - Hitelességellenőrzés (Fixity): biztosítja, hogy a tartalominformáció nem változott meg dokumentálatlanul
 - Hozzáférési jogok (Access Rights): rögzíti a megőrzés és a hozzáférés feltételeit

A Csomagolási információ (Packaging Information) teszi lehetővé, hogy az AIP egy logikai egységként azonosítható legyen.

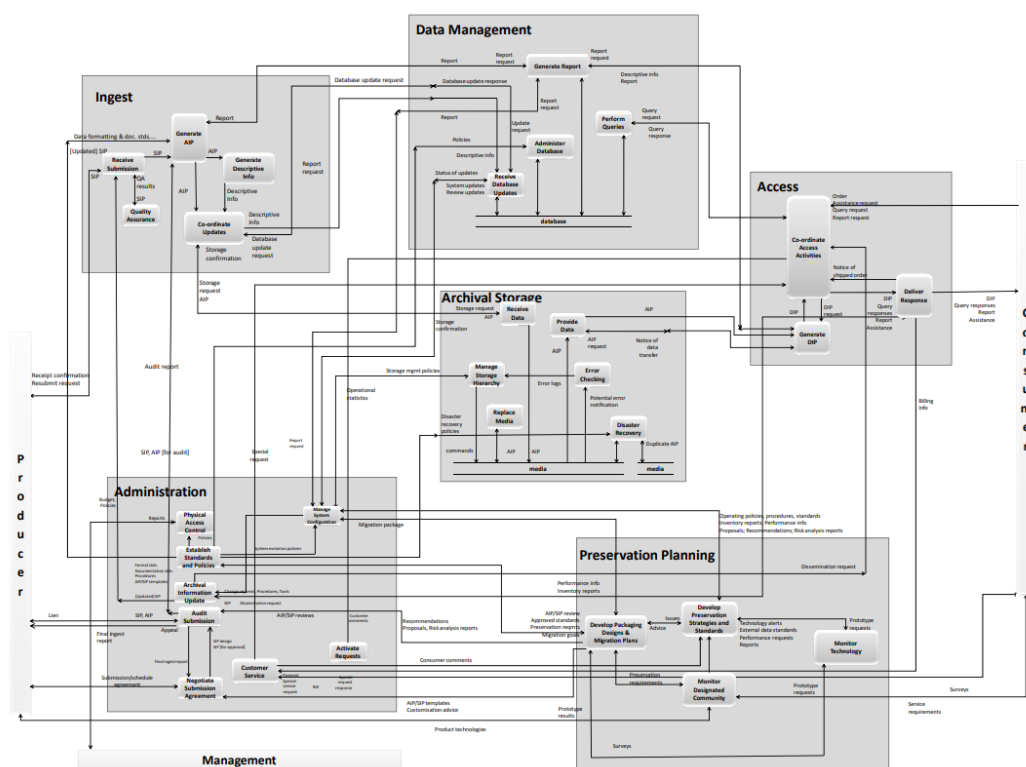
A Csomagleírás (Package Description) az AIP megtalálását szolgálja. Ez lehet például a Tartalominformáció és a Megőrzési leíró információ adataiból konvertált és válogatott Dublin Core sémájú rekord, amely segíti az archívumhasználót az archívum felfedezésében, a keresett anyag megtalálásában.

Funkcionális modell

Az OAIS referenciamodell hat funkcionális entitást határoz meg:

- Befogadás (Ingest): Feldolgozza a SIP-eket és átalakítja őket AIP-pá.
- Archív tár (Archival Storage): Tárolja, karbantartja és visszakereshetővé teszi az AIP-okat.

- **Adatmenedzsment (Data Management):** Kezeli a metaadatokat, biztosítva az adatok kereshetőségét és visszakereshetőségét.
- **Adminisztráció (Administration):** Irányítja az archívum napi működését.
- **Megőrzéstervezés (Preservation Planning):** Biztosítja, hogy az archívum alkalmazkodjon a technológiai változásokhoz. Figyelemmel kíséri a célközösséget (Designated Community) és a szélesebb technológiai környezetet, valamint meghatározza a követendő irányelveket és szabványokat.
- **Hozzáférés (Access):** Kommunikál az archívumhasználóval, és igény szerint létrehozza a DIP-eket.



4. ábra
Az OAIS funkcionális entitásai [11]

3.2 PREMIS

A PREMIS Data Dictionary for Preservation Metadata a digitális objektumok megőrzését támogató és hosszú távú használhatóságukat biztosító metaadatok nemzet-

közi szabványa. A nemzetközi szakértői csoport által kifejlesztett PREMIS-t világszerte alkalmazzák a digitális megőrzési projektekben és rendszerekben. A szabvány egy adatszótárból, egy XML-sémából és a támogató dokumentációból áll. [14]

A PREMIS adatmodell első két verziója öt, a harmadik – jelenleg utolsó, 2015-ben kiadott – verziója négy fő entitást foglal magában:

- **Objektumok (Objects):** A digitális megőrzési rendszerben tárolni és megőrizni kívánt elemek. A PREMIS négy különböző objektumtípust határoz meg: Intellektuális entitások (Intellectual Entities), Reprezentációk (Representations), Fájlok (Files) és Bitfolyamok (Bitstreams). Az objektumokhoz olyan metaadatok kapcsolódnak, mint például az azonosítók, az ellenőrzőösszegek, az objektum mérete, formátuma, egyéb technikai jellemzők, kapcsolatok más objektumokkal stb.
- **Események (Events):** Az objektumokat érintő műveletek, melyeknek naplózása alapvető fontosságú az információ életciklusának követése szempontjából. A PREMIS kontrollált szótárt ad az események lehetséges típusainak meghatározásához [15].
- **Jogok (Rights):** Az objektumokra vonatkozó jogi információkat tartalmazza, ideértve a szerzői jogokat és a digitális megőrzési folyamat során ellenőrizendő engedélyeket.
- **Szereplők¹ (Agents):** Az Eseményekkel és a Jogokkal kapcsolatos személyek, szervezetek, szoftverek és hardverek. Minden kapcsolathoz kötelezően rendelni kell szerepkört is (például szerző, szerzői jog tulajdonosa).

A PREMIS szabvány előnye, hogy lehetővé teszi az archívumi információk egységes, rendszerezett tárolását és könnyű adatcseréjét más rendszerekkel. A szabvány XML-alapú adatcserére épül, így egyszerűen integrálható más megőrzési és leíró szabványokkal, mint például a következő fejezetben leírt METS-szel. [17]

¹ Az Agents kifejezést Dancs Szabolcs Cselekvőknek fordítja [16], ezt megfelelőbb választásnak tartom az Ügynökök vagy Ágensek kifejezéseknél.

3.3 METS

A Library of Congress és a Digital Library Federation által felügyelt METS (Metadata Encoding and Transmission Standard) is egy XML-alapú szabvány, amelyet a digitális objektumok különféle metaadatainak – leíró, adminisztratív, technikai, strukturális és viselkedési adatok – kódolására és cseréjére használnak digitális könyvtárakban és archívumokban. A szabvány biztosítja, hogy az összetett digitális objektumok teljes struktúrája és a kapcsolódó információk egységes szerkezetben legyenek kezelhetők és cserélhetők a különböző rendszerek között, ezzel megkönnyítve az intézmények közötti együttműködést is. [18]

A METS dokumentum – legfeljebb – hét szakaszból áll:

- metsHdr: Fejléc (Header), a dokumentum létrehozásának és szerkesztésének metaadatai;
- dmdSec: Leíró metadatok (Descriptive Metadata), melyek szerepelhetnek belsőleg a dokumentumba ágyazva, de mutathatnak a dokumentumon kívülre is, például egy OPAC MARC rekordjára vagy egy online EAD rekordra;
- amdSec: Adminisztratív metadatok (Administrative Metadata), beleértve a fájlok eredetét és szerzői jogi információkat;
- fileSec: Fájlcsoportok (File Groups), a digitális objektumhoz tartozó fájlokat listázza csoportosítva;
- structMap: Szerkezeti térkép (Structural Map), a dokumentum hierarchikus szerkezetét írja le;
- structLink: Szerkezeti kapcsolatok (Structural Links), hiperhivatkozásokat rögzít a szerkezeten belüli csomópontok között, különösen hasznos weblapok archiválásakor;
- behaviorSec: Viselkedés (Behavior), alkalmazások, futtatható kódok, melyek szükségesek lehetnek a digitális objektum megjelenítéséhez vagy konvertálásához.

Sok dokumentumtípushoz (pl. könyvek, kották, hangfelvételek, videófelvevételek, fotók) már léteznek kidolgozott METS profilok. Egy példa a brit nemzeti könyvtár által kidolgozott BLAP-S (British Library Application Profile – Sound) profil,

amely támogatja a szegmentált hangfájlok adatainak rögzítését. Ez a profil alkalmas olyan több órás interjúk és egyéb hosszú hangfelvételek kezelésére, amelyeket érdemes kisebb részekre bontva adatolni és tárolni, de nem szabad elveszíteni az egyes szegmensek közötti kapcsolatokat megadó információt. [19]

3.4 EAD

Az EAD (Encoded Archival Description) is egy XML formátumot használó nemzetközi szabvány, amely elsősorban a levéltári anyagok leírását és megosztását szolgálja. Célja, hogy egységesítse a levéltári leírásokat, elősegítve azok cseréjét és hozzáférhetőségét különböző intézmények között. Az EAD támogatja azokat a hierarchikus leírási struktúrákat, amelyek a levéltári anyagok összetett szerkezetét tükrözik. A szabványt a Society of American Archivists (SAA) és a Library of Congress közösen fejlesztette ki és tartja karban. [20]

Az EAD különösen hasznos a levéltári adatok online megjelenítésében, mivel a strukturált XML-formátum könnyen átalakítható weboldalra vagy PDF-re, és támogatja a szofisztikált keresési funkciókat is. A legfrissebb verzió, az EAD3 újabb elemekkel bővítette az EAD 2002 szabványt, bővítette az archívumi rekordok származásának és kezelésének dokumentálási lehetőségeit. [21]

3.5 ISAD(G)

Az ISAD(G) – General International Standard Archival Description az általános levéltári leírás nemzetközi szabványa. Az ISAD(G) a leírások hierarchikus rendszerét támogatja, lehetővé teszi a levéltári anyagok különböző szinteken – például fondok, állagok, sorozatok, tételek, ügyiratok, iratdarabok – való feltárását.

A szabványban a leírás elemei hét adatcsoportban adhatók meg:

- Azonosítási adatcsoport
- Kontextusra vonatkozó adatcsoport
- Tartalomra és szerkezetre vonatkozó adatcsoport
- Hozzáférésére és használatra vonatkozó adatcsoport
- Kapcsolódó anyagokra vonatkozó adatcsoport
- Megjegyzések adatcsoportja
- Ellenőrző adatcsoport

A szabvány 26 elemet tartalmaz, közülük 6 kötelezően megadandó a nemzetközi levéltári adatcseréhez:

- jelzet
- cím
- iratképző
- idő(kör)
- terjedelem
- leírás szintje

A szabvány 1999-ben elfogadott második kiadásához 2009-ben magyar fordítás készült. [22]

3.6 Dublin Core

A Dublin Core Metadata Initiative (DCMI) által támogatott szabvány célja, hogy megkönnyítse az adatcserét és javítsa a kereshetőséget. Az egyik legelső leíró metaadat-szabvány, amely – egyszerűsége és általánosíthatósága miatt is – nagyon széles körben elterjedt. A szabvány két változatban érhető el: a „Simple Dublin Core” 15 alaptulajdonságot, attribútumot tartalmaz, míg a „Qualified Dublin Core” további lehetőséget biztosít a metaadatok pontosabb meghatározásához, például a jogok és a hozzáférési korlátozások rögzítéséhez.

A 15 alapvető attribútum:

- Cím (Title)
- Létrehozó (Creator)
- Tárgy (Subject)
- Leírás (Description)
- Kiadó (Publisher)
- Közreműködő (Contributor)
- Dátum (Date)
- Típus (Type)
- Formátum (Format)
- Azonosító (Identifier)
- Forrás (Source)

- Nyelv (Language)
- Kapcsolat (Relation)
- Térbeli vagy időbeli lefedettség (Coverage)
- Jogok (Rights)

A Dublin Core alapú metaadat-struktúra RDF/XML formátumban is megadható, ekkor a szemantikus web RDF (Resource Description Framework) szintaxisát használjuk a strukturált adatok és kapcsolataik kifejezésére. Ez a formátum biztosítja, hogy az attribútumok gépi feldolgozással azonosíthatók legyenek, ezáltal a leírt adatok összekapcsolhatók legyenek más forrásokban található adatokkal. Az RDF alapú Dublin Core leírások sokkal gazdagabb keresési, adatkezelési és adatelemzési lehetőségeket biztosítanak, mint a hagyományos adatstruktúrák. [23]

3.7 OAI-PMH

Az OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) egy HTTP protokoll, amelyet a metaadatok időzített, automatikus és szabványos gépi begyűjtésére (harvesting) fejlesztettek ki. Ennek a protokollnak a segítségével gyűjtik be az aggregátorok az elosztott repozitóriumokból a metaadatokat annak érdekében, hogy valamilyen szolgáltatást, például közös keresési vagy adatbányászati lehetőséget biztosítsanak. Két rendszer közötti adatszinkronizációra is használható.

Az OAI-PMH főként XML formátumú metaadatokat használ, és leggyakrabban a Dublin Core elemekre épít. A protokoll két fő szereplőt különböztet meg:

- Adatszolgáltatók (Data Providers), amelyek a metaadatokat elérhetővé teszik az OAI-PMH interfészen keresztül.
- Adatgyűjtők (Service Providers), amelyek lekérdezik és összegyűjtik ezeket a metaadatokat a különböző adatszolgáltatóktól az OAI-PMH interfészen keresztül.

Az OAI-PMH hat protokollparancsot – műveletet vagy „igét” (verb) – definiál, melyeknek segítségével az adatgyűjtő lekérheti az adatszolgáltatótól a metaadatokat:

- Identify: információ az adatszolgáltatóról;
- ListMetadataFormats: metaadat-formátumok listája;
- ListSets: gyűjtemények listázása;

- GetRecord: egy konkrét rekord lekérdezése azonosító alapján;
- ListIdentifiers: rekordazonosítók listázása;
- ListRecords: rekordok teljes vagy szűrt listája. [24]

DSpace OAI-PMH Data Provider

Identify

Sets

Records

Identifiers

Metadata Formats

Response Date 2024-11-01 20:05:50

Repository Information

Repository Name	ELTE Digitális Intézményi Tudástár
E-Mail Contact	admin@edit.elte.hu
Repository Identifier	
Sample identifier	
Protocol Version	2.0
Earliest Registered Date	2013-09-27 10:47:38
Date Granularity	YYYY-MM-DD hh:mm:ss
Deletion Mode	transient

5. ábra

Az ELTE Digitális Intézményi Tudástár OAI-PMH lekérdezése az Identify protokollparanccsal

Az 5. ábra a <https://edit.elte.hu/oai/request?verb=Identify> lekérdezésre adott választ mutatja ember számára könnyen olvasható formában, de mögötte a gépek számára jól értelmezhető xml formátumú eredmény van.

A https://edit.elte.hu/oai/request?verb=ListRecords&metadataPrefix=oai_dc kérésre 2024. november 1-én 14157 rekordot kapunk, melyet a set=com_10831_4 kiegészítéssel leszűrhetünk a Szakdolgozatok gyűjteményére. Remélhetőleg hamarosan ennek a dolgozatnak a metaadatait is lekérdezhethetjük a GetRecord paranccsal.

Az interoperabilitás, vagyis együttműködési képesség nagyon fontos szempont az archívum szoftvereinek kiválasztásakor. A webes integrációban egyre elterjedtebb JSON formátumú API kapcsolatok lehetősége mellett az OAI-PMH interfész biztosítását is érdemes vizsgálni.

4. NEMZETKÖZI KUTATÁSOK ÉS TAPASZTALATOK

4.1 E-ARK program

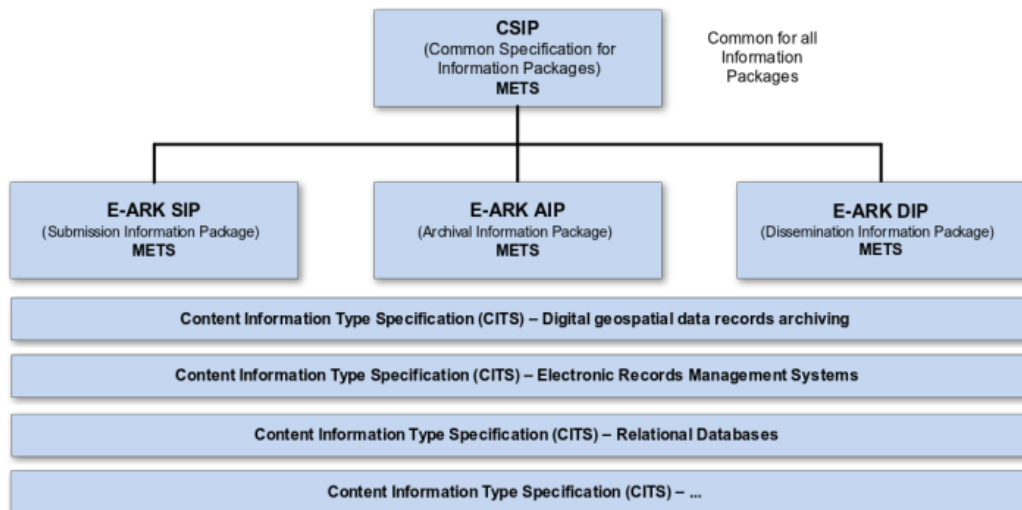
Az E-ARK (European Archival Records and Knowledge Preservation) program egy 2014-ben elindított európai uniós projekt, amelynek célja a digitális archiválási módszerek és technológiák fejlesztése az archívumi, elsősorban a levéltári megőrzési gyakorlatok harmonizálása és országhatárokon átnyúló integrációja érdekében. Az E-ARK módszerei, eszközei és infrastruktúrája nemcsak a nemzeti levéltárak, hanem a regionális, helyi archívumok, piaci vállalatok, felsőoktatási, tudományos és kulturális intézmények számára is hasznosak. A program különböző elnevezésekkel már több ciklust megélt, jelenleg eArchiving Initiative néven fut. Az elkészült specifikációkat a Digital Information LifeCycle Interoperability Standards Board (DILCIS Board) gondozza és publikálja.

Az E-ARK szabványai és eszközei az OAIS referenciamodellen alapulnak. Az információs csomagok leírásának fő szabványa a METS, az információs csomagok megőrzésének fő szabványa a PREMIS. Az E-ARK program egyik célja az volt, hogy a referenciamodellben szereplő SIP, AIP és DIP csomagokra vonatkozóan technikai implementációtól független, részletes specifikációt készítsen. Ezt úgy oldották meg, hogy a 3 csomag mindegyikére vonatkozó közös részt kiemelték a Common Specification for Information Packages (CSIP) specifikációba, és csak a különbségeket definiálja az E-ARK SIP, E-ARK AIP és E-ARK DIP. Például az E-ARK SIP tartalmazza a benyújtási megállapodás (submission agreement) és a benyújtókkal (producers) kapcsolatos releváns információk rögzítésének követelményeit, az E-ARK DIP pedig az archivált adatok újrafelhasználásához szükséges környezet leírásának követelményét. [25]

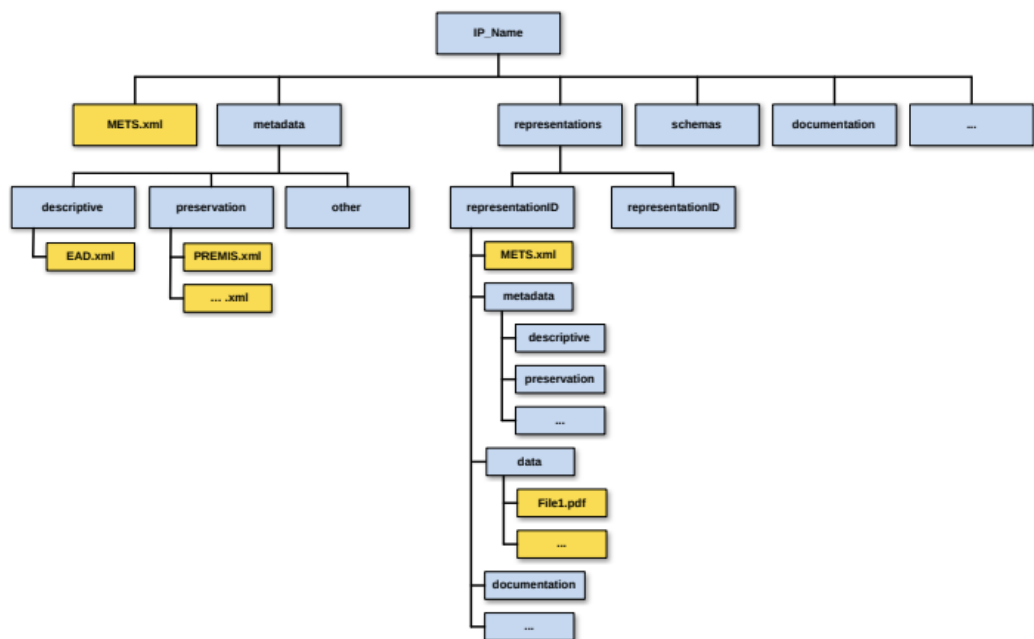
A Content Information Type Specification (CITS) célja, hogy definiálja, hogy speciális tartalomtípusok (pl. térinformatikai adatok, dokumentumok, adatbázisok, egészségügyi adatok stb.) esetén hogyan kell rögzíteni az adatokat és a metaadatokat.

Az 7. ábra látható a CSIP mappa- és fájlstruktúrája, amely az általános OAIS modellt a csomagszerkezetre vonatkozó konkrét ajánlással egészíti ki. Minimális követelmény, hogy kell lenni egy gyökérmappának, ami tartalmaz egy METS.xml

fájlt, egy metadata és egy representations mappát. A specifikáció részletei megtalálhatók a DILCIS Board weboldalának Specifications menüpontja alatt, közérthető nyelven, olvasmányosan.



6. ábra
Az E-ARK specifikáció vázlata [25]



7. ábra
A CSIP mappastruktúrája [25]

Az OAIS referenciamodellt implementálni kívánó archívumoknak és a digitális archiválás iránt érdeklődő szakembereknek ad iránymutatást az eArchiving Reference Architecture (Digitális archiválás referenciaarchitektúrája). A referenciaarchitektúra kidolgozását végző munkacsoport az alábbi, elméleti alapot biztosító elvek, általános szabályok alapján dolgozik.

Üzleti alapelvek

1. Az archiválás célja a felhasználás.
2. Az archiválás az információ megőrzése technológiák segítségével.
3. Az archiválás biztosítja az információ megbízhatóságát.
4. Az archiválás az információ és folyamatok rendszerezése és tervezése.
5. Az archiválás pragmatikus, hatékony, rugalmas és fenntartható.
6. Az archiválás az információkezelés szerves része.
7. Az archiválás figyelembe veszi a jogokat, engedélyeket és korlátozásokat.

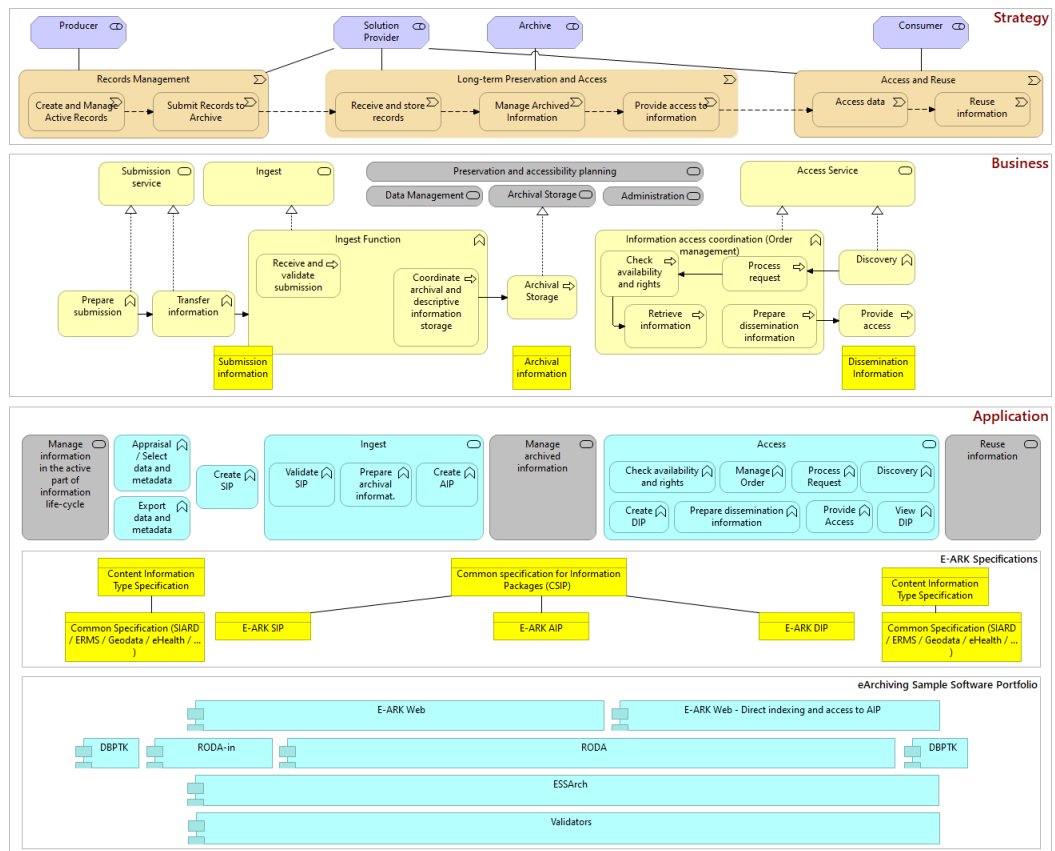
Információs alapelvek

8. Az archivált adatokhoz megfelelő metaadatok és dokumentáció kapcsolódik.
9. Az archivált információ hosszú távon visszakereshető.
10. Az archivált információk közös szótárakat és definíciókat használnak.
11. Az archivált információk gépi úton is felhasználhatók.

Alkalmazási elvek

12. Az archiváló rendszereknek tervezniük és kezelniük kell saját elavulásukat.
13. Az archiválás tetszőleges informatikai rendszerben megvalósítható.
14. Az archiválási rendszerek és összetevőik interoperábilisak. [26]

A modell az ArchiMate üzleti modellező nyelv [27] szabályaival kialakított diagramok segítségével mutatja be az archiválási folyamatok összefüggéseit négy rétegben (motiváció, stratégia, üzlet, alkalmazás). Szemléltetésként egy kiragadott példa látható a 8. ábra.



8. ábra

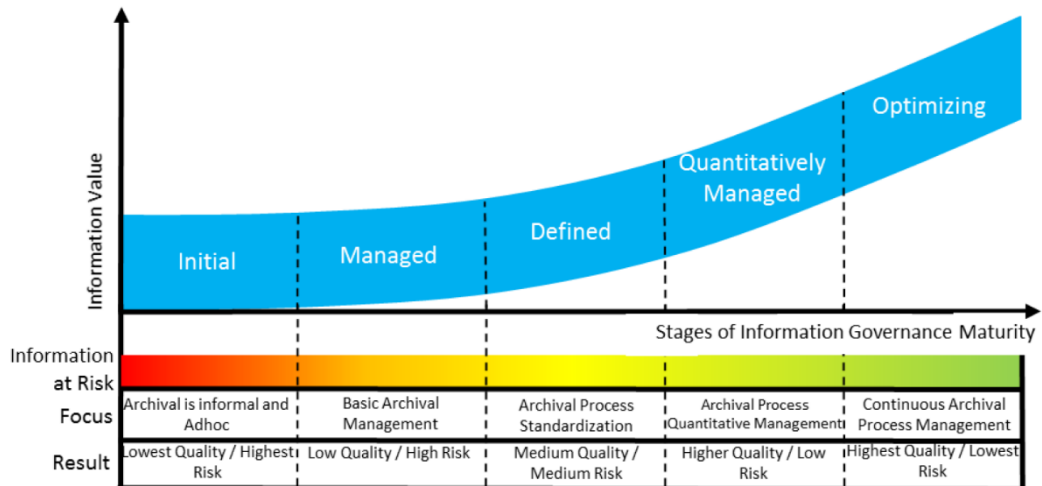
Az eArchiving Reference Architecture 2.0 központi modelljének egyik áttekintő nézete [28]

Az eArchiving Reference Architecture modellben nagy hangsúlyt fektetnek az OAIS-ből hiányzó Pre-Ingest folyamat kidolgozására is.

Az eArchiving Reference Architecture 2.0 változatában megjelenik az „Archiving by Design” megközelítés modellezése is. A magyarra *tudatosan tervezett archiválásként* fordítható koncepció lényege, hogy az információ előállításától kezdve az információ teljes életciklusában folyamatosan szem előtt kell tartani a fenntartható hozzáférhetőség szempontjait, melyek a kereshetőség, elérhetőség, olvashatóság, értelmezhetőség, megbízhatóság és a jövőállóság. [29]

Egy szervezet digitális archiválási folyamatainak javítását, az egyre szervezettebb és rendszerezettebb megőrzési képességek kiépítését segíti az eArchiving Capability Maturity Model (eACMM), mely az üzleti világban és az akadémiai körökben is elterjedten használt érettségi modellek alkalmazása erre a speciális témára. A digitális archiválási képességek érettségi modellje önértékelési eszközt ad a szerve-

zetek kezébe, segíti az erősségek és a fejlesztést igénylő gyenge pontok azonosítását, priorizálja a szükséges lépéseket a magasabb szint eléréséhez. A képességek korrelálnak az OAIS-ból ismert funkcionális entitásokkal (Pre-Ingest, Ingest, Archival Storage and Preservation, Data Management, Access, Preservation and Accessibility Planning), az egyes képességekhez rendelt szintek a 9. ábra grafikonján láthatók. [30]



9. ábra
Az eACMM érettségi táblája [30]

Az eACMM szerinti elemzéshez Excel-formátumú űrlap is készült, melynek segítségével bármely szervezet önértékelést végezhet. A Folklór Archívum is használhatja majd az új rendszer kiépítésének folyamatában, a prioritások meghatározásában.

4.2 ISA² program

Az ISA² („Interoperability Solutions for European Public Administrations”) program az európai közszolgáltatások átjárhatóságát, együttműködését, interoperabilitását vizsgáló és támogató projekt második üteme volt, melyet az Európai Bizottság finanszírozott 2016 és 2020 között [31]. A program része volt az ISA² Action 2017.01 azonosítójú „Standard-Based Archival Data Management, Exchange and Publication”, magyarul „Szabványokon alapuló archiválási adatkezelés, adatcsere

és publikálás” megnevezésű szerteágazó kutatás, melynek eredményeit egy 315 oldalas jelentésben tették közzé [32]. A kutatásban a Magyar Nemzeti Levéltár is részt vett a DLM Forum² tagjaként.

A kutatásban az általánosan érvényes következtetések levonása mellett az is megjelent célként, hogy megvizsgálják az ajánlások megvalósíthatóságát a Historical Archives Service (HAS) új archívumkezelő rendszerének bevezetésével kapcsolatban. A tanulmány részletesen elemzi az üzleti folyamatokat, kiterjesztve az OAIS modellt úgy, hogy támogassa az archiválási tevékenységek teljes spektrumát, nagy hangsúlyt fektetve az interoperabilitásra és az adatcsere szükségességére. Emellett összegyűjti és leírja az általánosan elfogadott archívumi szabványokat, megvizsgálja a Linked Open Data alkalmazási lehetőségeit, alaposan elemzi az elterjedtebb szoftvereket, és megállapítja, hogy a gyűjteménykezelő és megőrzési rendszerek kombinált megvalósítása lefedheti a memóriaintézmények belső üzleti folyamatainak nagy részét.



10. ábra

„Figure 32: Software stack” [32]

Kifejezetten a HAS igényeit és folyamatait megvizsgálva, háromféle szoftvercsomagot ajánl a tanulmány (10. ábra):

1. ArchivesSpace (nyílt forráskódú) és Preservica (zárt forráskódú)
2. AtoM (nyílt forráskódú) és Archivematica (nyílt forráskódú)
3. Archeevo (zárt forráskódú) és RODA (nyílt forráskódú)

A következő fejezetben megvizsgálunk néhányat az ajánlott nyílt forráskódú szoftverek közül.

² A DLM a Document Lifecycle Management rövidítése. A DLM Forum egy európai együttműködési platform, amely kormányzati és magánszektorbeli szakértőket fog össze. [33]

5. NYÍLT FORRÁSKÓDÚ ÉS SZABVÁNYOS MEGOLDÁSOKAT KÍNÁLÓ SZOFTVEREK

5.1 A nyílt forráskódú szoftverek előnyei és hátrányai

A gyűjteménykezelő és a hosszú távú megőrzést biztosító szoftverek kiválasztásánál az egyik alapvető szempontunk a költséghatékonyság a rendszer bevezetése és fenntartása során. Mivel a nyílt forráskódú szoftverek sok esetben ingyenesen hozzáférhetők és ez jelentős megtakarítást jelenthet a licencdíjas megoldásokhoz képest, ezért elsősorban ezeket a lehetőségeket vizsgáljuk az új rendszer tervezése folyamán.

A nyílt forráskódú szoftverek egyik legnagyobb előnye, hogy átlátható és nyilvános a fejlesztési folyamatuk. A felhasználók megismerhetik és ellenőrizhetik a forráskódot, a program működését, aktívan részt vehetnek a hibajavításban, továbbfejlesztésben, új funkciók létrehozásában. Ez a közösségi fejlesztés hozzájárul a szoftver minőségének emeléséhez. További előny, hogy ezek a szoftverek az egyedi igényekhez is rugalmasan hozzáigazíthatóak, testre szabhatóak. A választás mellett szól az is, hogy a felhasználó nincs egy adott gyártóhoz, szállítóhoz kötve.

A nyílt forráskódú szoftvereknek természetesen vannak hátrányai is. Gyakran speciális technikai ismereteket követel a telepítés és a karbantartás, ami egy olyan intézmény számára, amelynek nincs erre dedikált szakértő alkalmazottja, problémát jelent. Bár sok esetben léteznek fórumok és elérhetők a dokumentációk, nem mindig áll rendelkezésre hivatalos támogatás, így a problémák megoldása hosszabb időt vehet igénybe. Biztonsági kockázatokat is rejthet a nyílt forráskód, mert rosszindulatú szereplők visszaélhetnek az esetleges hibákkal, kihasználhatják a könnyebben megismerhető sebezhetőséget. A nyílt forráskódú projektek gyors fejlődése miatt a frissítések követése, a naprakészen tartás is kihívást jelenthet.

5.2 Gyűjteménykezelő rendszerek

Az angolul Archive Management Systems vagy Collection Management Systems megnevezésű gyűjteményi adatkezelő rendszerek közül két, széles körben használt és aktív támogatói közösséggel rendelkező nyílt forráskódú megoldást vizsgáltunk.

Ezek az ArchivesSpace [34] és az Access to Memory (AtoM) [35] rendszerek, melyeket telepítettünk is tesztelés céljából.

A saját tapasztalataink mellett más intézmények szofterválasztási szempontjait és értékeléseit is meg akartuk ismerni. A Society of American Archives kérdőíves felmérést készít az amerikai archívumok által használt CMS (Collection Management System) és DAMS (Digital Assets Management System) rendszerekről, melyben a kiválasztás szempontjaira és a választott szoftverrel kapcsolatos tapasztalatokra is rákérdez, és a válaszokat közzéteszi. [36] A University of York AtoM bevezetéséről [37] és a University of Edinburgh ArchivesSpace bevezetéséről [38] írt blogposztok is sok hasznos információt tartalmaznak.

	ArchivesSpace	AtoM
Programozás nyelve	Ruby	PHP
Licensz	GNU General Public License v3.0	GNU Affero General Public License v3.0
Népszerűség	4,7 (növekvő)	3,9 (csökkenő)
Aktivitás	9,7 (növekvő)	8,3 (csökkenő)
Csillagok száma (Stars)	351	248
Követők száma (Watchers)	73	43
Feljesztési elágaztatások száma (Forks)	228	127
Kiadási ciklus (Release cycle)	55 nap	34 nap
Utolsó verzió	2020.11.11.	2020.10.21.
Utolsó frissítés (Last commit)	2024.10.31.	2024.10.31.

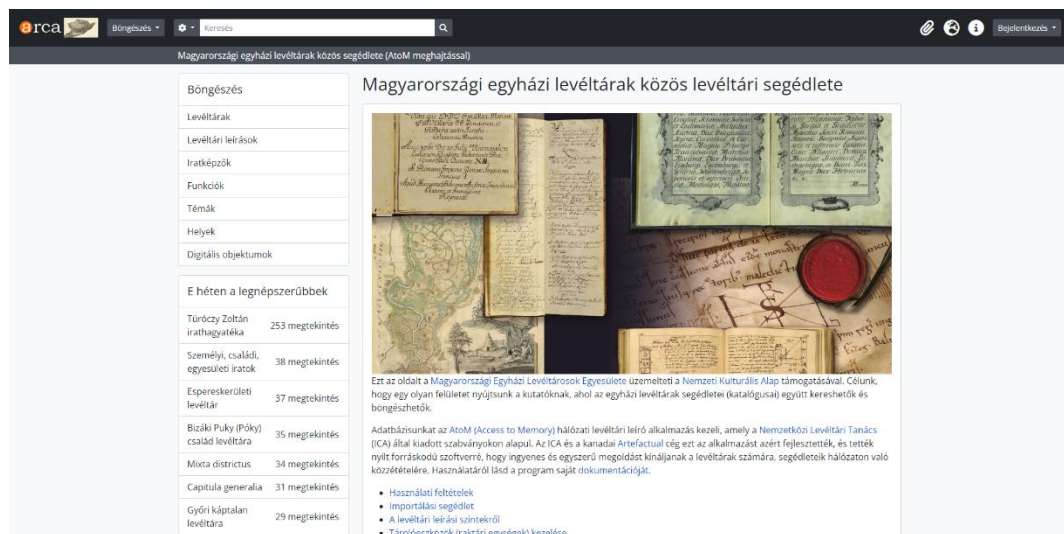
1. táblázat

Az ArchivesSpace és az AtoM összehasonlítása Github-adatok alapján [39] [40] [41]

A forráskódokat tároló Github-repozitóriumok 2024. november 2-i adatai alapján készült a két szoftver összehasonlítását megjelenítő 1. táblázat. Látható, hogy mindkét szoftver használata széles körben elterjedt és aktív fejlesztői támogatással rendelkezik.

Összevetve az ArchivesSpace és az AtoM legfontosabb entitásait megállapítottuk, hogy nagyjából megfeleltethetőek egymásnak: Repository – Archival institution; Accession – Accession; Resource – Archival description; Digital object – Digital object; Agent – Actor; Event – Event stb. Mindkettő adatbázis-szerkezete alkalmas a Folklór Archívum gyűjteményének és folyamatainak leírására.

Mindkét rendszer jól dokumentált integrációs lehetőségekkel, API-val, OAI-PMH interfésszel és export/import funkciókkal rendelkezik.



11. ábra
Egy magyar példa az AtoM alkalmazására:
A Magyarországi egyházi levéltárak közös segédlete [41]

A szoftver kiválasztásához más szempontokat is figyelembe kell venni. A két rendszer közötti főbb eltéréseket foglalja össze a 2. táblázat. Az AtoM választása mellett szól, hogy közvetlenül kezeli a digitális objektumokat, így a médiafájlokhoz való publikus hozzáféréshez nem szükséges más szoftvert integrálni.

A megfelelő gyűjteménykezelő szoftver kiválasztási folyamata még zajlik. A tervezést végző projektcsapat folytatja a tesztelést. Jelenleg az importálási lehetőségeket vizsgáljuk, illetve azt próbáljuk felmérni, hogy mekkora külső fejlesztői kapacitást igényelne az egyes rendszerek bevezetése és támogatása.

	ArchivesSpace	AtoM
Írányadó metaadat-szabvány	EAD	ISAD(G)
Nyelvi támogatás	angol, de kiegészíthető saját nyelvi verzióval	többnyelvű, van magyar nyelvi változata
Elterjedtség	leginkább az Egyesült Államokban használják, de vannak európai példák is	nemzetközi, néhány magyar példa is (11. ábra)
Fizikai objektumok tárolási helyének kezelése	robosztus	korlátozott
Digitális objektumok kezelése	a fájlokat más rendszerbe kell betölteni, itt csak az elérési útvonal adható meg	a fájlok betölthetők a rendszerbe
Támogatás	széles, főként amerikai közösségi támogatás, tagdíj-alapú hozzáférés bizonyos szolgáltatásokhoz	széles nemzetközi közösségi támogatás, ingyenes hozzáférés minden támogatói anyaghoz

2. táblázat
ArchivesSpace és AtoM összehasonlítása saját szempontok szerint

5.3 Hosszú távú megőrzésre szolgáló rendszerek

Az angolul Preservation Systems megnevezésű, hosszú távú megőrzésre szolgáló rendszerek közül is két nyílt forráskódú, OAIS modell alapján működő, archívumi szabványokat (pl. PREMIS, METS) követő, webalapú rendszert, az Archivematica [43] és a RODA Community Edition [44] szoftvereket vizsgáltuk. Ezeket egyelőre nem telepítettük, csak a nyilvánosan elérhető demóoldalakon teszteltük. [45], [46]

A RODA Community Edition demójával már megismerkedtem az adatbázis-archiválásról szóló szakmai gyakorlatom során [47]. A SIP csomag könnyen elkészíthető a KEEPS RODA-in eszközével Bagit vagy E-ARK formátumban [48].

A 3. táblázat összehasonlító adataiból látható, hogy mindkét nyílt forráskódú szoftver közösségi támogatása és fejlesztése jelentős. Az Archivemática szélesebb körben, globálisan elterjedt. A RODA elsősorban európai intézmények körében népszerű, különösen Portugáliában, ahol a fejlesztői bázisa is található.

	Archivemática	RODA CE
Programozás nyelve	Python	JAVA
Licensz	AGPL-3.0	LGPL-3.0
Csillagok száma (Stars)	428	74
Követők száma (Watchers)	44	13
Feljesztési elágaztatások száma (Forks)	103	48
Hozzájáruló fejlesztők száma	44	22
Utolsó verzió	2024.05.16.	2024.09.05.
Utolsó frissítés (Last commit)	2024.10.29.	2024.09.05.

3. táblázat
Az Archivemática és a RODA Community Edition összehasonlítása
Github-adatok alapján [49], [50]

Még nem végeztünk elég vizsgálatot ahhoz, hogy dönteni tudjunk a két rendszer között. Az Archivemática választása mellett szól, hogy az előző alfejezetben tárgyalt gyűjteménykezelő rendszerek mindegyikéhez van már kidolgozott integrációs megoldása, melyet sok intézmény használ. A RODA mellett szól, hogy a felülete intuitív, felhasználóbarát, figyelembe veszi az európai szabályozásokat és elvárásokat.

Látható a fentiekből, hogy a Folklór Archívum új gyűjteménykezelési és megőrzési rendszerének specifikációja még nem befejezett, a tervezési és tesztelési munkálatok java még hátra van.

6. EREDMÉNYEK ÉS TERVEK

A Folklór Archívum gyűjteménykezelési és megőrzési rendszerének megújításáról szóló projekt 2024 őszén kezdődő tervezési szakaszának legelső ütemében az volt a célunk, hogy felmérjük a követendő szabványokat, a hazai és nemzetközi gyakorlatot, a hasonló kihívásokkal rendelkező archívumok tapasztalatait és megoldásait. A szakdolgozatom célközönsége – lásd még „designated community” – nemcsak a felkért bíráló, hanem a Hagyományok Házában dolgozó munkatársaim, az általam vezetett projektcsapat tagjai is.

A jelenleg zajló második ütemben az alkalmazható szoftvereket kutatjuk és teszteljük. Ennek a projektszakasznak még az elején járunk, de az eddigi eredmények biztatóak. Feltehetően nem lesz egyedi szoftverfejlesztésre szükség, mert nyílt forráskódú alkalmazások integrációjával felépíthető az új rendszer. Az éles környezet kialakítása, a telepítés, a konfigurálás, a testreszabás, az integráció és a rendszerek üzemeltetése viszont számottevő erőforrást igényel majd.

Rövid távon az a célunk, hogy – a hosszú távú megőrzés szempontjait és a későbbi integrációs követelményeket is figyelembe véve – kiválasszuk a gyűjteménykezelő szoftvert, megtaláljuk az éles működtetéshez a legmegfelelőbb környezetet és IT támogatást, és kidolgozzuk az adatok importálásának folyamatát. Ha ez az alrendszer már jól működik és használatát bevezettük az intézményben, akkor lépünk tovább a szintén alapvető fontosságú megőrzési rendszer megvalósítására.

7. IRODALOMJEGYZÉK

- [1] „A Hagymányok Háza Szervezeti és Működési Szabályzata”. Hagymányok Háza, 2023. Elérés: 2024. november 3. [Online]. Elérhető: https://hagyomanyokhaza.hu/sites/default/files/2023-02/HH_SZMSZ_matrix_2023.pdf
- [2] „Folklóradatbázis”. Elérés: 2024. szeptember 20. [Online]. Elérhető: <https://beta.folkloradatbazis.hu/>
- [3] Halász P., „Hatvan éves a Magyar Művelődési Intézet”, *Szín*, 11. évf., 6. sz., 22-26. o., 2006.
- [4] Sebő F., „A Népművészeti Intézettől a Hagymányok Házáig”, *Szín*, 12. évf., 6. sz., 19-21. o., 2007.
- [5] Földiák A., „A Magyar Művelődési Intézet ötven évének öt korszakáról”, *Szín*, 1. évf., 2. sz., 2-7. o., 1996.
- [6] Grozdits K., „Csak emlék a Néptáncosok Szakmai Háza? Beszélgetés Stoller Antallal”, *Folkmagazin*, 4. sz., 8-11. o., 2016.
- [7] Pávai I., „Hagymányörző Műhely – Martin Médiatár a Corvin téren”, *Folkmagazin*, 6. évf., 2. sz., 9 o., 1999.
- [8] „Gyűjtőkori és Állományvédelmi Szabályzat”. Hagymányok Háza, 2021. Elérés: 2024. október 19. [Online]. Elérhető: <https://hagyomanyokhaza.hu/sites/default/files/2024-04/Gy%C5%B1jt%C5%91k%C3%B6ri%20Szab%C3%A1lyzat.pdf>
- [9] „Gyűjtemény- és Raktárkezelési Szabályzat”. Hagymányok Háza, 2021. Elérés: 2024. október 19. [Online]. Elérhető: https://hagyomanyokhaza.hu/sites/default/files/2021-02/Gy%C5%B1jtem%C3%A9ny-%20%C3%A9s%20Rakt%C3%A1rkezel%C3%A9si%20Szab%C3%A1lyzat_20210204_1.pdf
- [10] „Múzeumi spektrum”. Magyar Nemzeti Múzeum Országos Muzeológiai Módszertani és Információs Központ, 2020. Elérés: 2024. október 19. [Online]. Elérhető: <https://muzeumispektrum.hu/index.php/Folyamatok>
- [11] „Reference Model for an Open Archival Information System (OAIS)”. CCSDS Secretariat, 2012. Elérés: 2024. november 1. [Online]. Elérhető: <https://public.ccsds.org/pubs/650x0m2.pdf>
- [12] Bódog A., „A nyílt archívumi információs rendszer (OAIS) szabványának honosítása”, OSZK. Elérés: 2024. október 12. [Online]. Elérhető: https://nemzetikonyvtar.blog.hu/2022/04/11/a_nyilt_archivumi_informacios_rendszer_oais_szabvanyanak_honositasa
- [13] „MSZ ISO 14721:2022 szabvány adatai”. Elérés: 2024. október 12. [Online]. Elérhető: <https://ugyintezes.mszt.hu/webaruhaz/szabvany-adatok?standard=140845>
- [14] „PREMIS: Preservation Metadata Maintenance Activity | Library of Congress”. Elérés: 2024. október 23. [Online]. Elérhető: <https://www.loc.gov/standards/premis/>
- [15] „Event Type - LC Linked Data Service: Authorities and Vocabularies (Library of Congress)”. Elérés: 2024. november 1. [Online]. Elérhető: <https://id.loc.gov/vocabulary/preservation/eventType.html>
- [16] Dancs S., „A hosszú távú megőrzés szabványos útja (és a Rosetta rendszer példája)”, *Tudományos és Műszaki Tájékoztatás*, 63 évf., 1 sz., 27-37. o., 2016.

- [17] P. Caplan, „Understanding PREMIS”. Library of Congress Network Development and MARC Standards Office, 2021. Elérés: 2024. november 1. [Online]. Elérhető: https://www.loc.gov/standards/premis/understandingPREMIS_english_2021.pdf
- [18] „Metadata Encoding and Transmission Standard (METS) | Library of Congress”. Elérés: 2024. november 1. [Online]. Elérhető: <https://www.loc.gov/standards/mets/>
- [19] „Digital Preservation Case Notes. Archival Sounds”. The British Library, 2010. Elérés: 2024. november 1. [Online]. Elérhető: <https://www.dpconline.org/docs/miscellaneous/advice/474-casenoteasr2/file>
- [20] „EAD: Encoded Archival Description | Library of Congress”. Elérés: 2024. november 1. [Online]. Elérhető: <https://www.loc.gov/ead/>
- [21] „Encoded Archival Description (EAD)”. Society of American Archivists, 2015. Elérés: 2024. november 1. [Online]. Elérhető: <https://www2.archivists.org/groups/technical-subcommittee-on-encoded-archival-standards-ts-eas/encoded-archival-description-ead>
- [22] *ISAD(G): General international standard archival description; adopted by the Committee on Descriptive Standards, Stockholm, Sweden, 19-22 Sept. 1999*, 2. Ed. in ICA Standards. Ottawa: International Council of Archives, 2000.
- [23] „DCMI Metadata Terms”, DCMI. Elérés: 2024. november 1. [Online]. Elérhető: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- [24] C. Lagoze, H. Van de Sompel, M. Nelson, és S. Warner, „Open Archives Initiative - Protocol for Metadata Harvesting - v.2.0”. Elérés: 2024. november 1. [Online]. Elérhető: <https://www.openarchives.org/OAI/openarchivesprotocol.html>
- [25] „E-ARK CSIP Common Specification v2.2.0”. DILCIS Board, 2024. [Online]. Elérhető: <https://earkcsip.dilcis.eu/>
- [26] Alföldi I., „E-ARK Reference Architecture - Principles v1.0”. DLM Forum, 2020. [Online]. Elérhető: <https://kc.dlmforum.eu/earchiving-ra20/download/E-ARK%20Reference%20Architecture%20-%20Principles%20v1.0.docx>
- [27] „The ArchiMate® Enterprise Architecture Modeling Language”. Elérés: 2024. október 29. [Online]. Elérhető: <https://www.opengroup.org/archimate-forum/archimate-overview>
- [28] „eArchiving RA | Core Views | Application | Overview”, eArchiving Reference Architecture v2.0. Elérés: 2024. november 4. [Online]. Elérhető: <https://kc.dlmforum.eu/earchiving-ra20/?view=5f44dd86-9620-440f-a5a7-4753c852854c>
- [29] „Archiving by Design”. European Archives Group, 2023. Elérés: 2024. október 29. [Online]. Elérhető: https://commission.europa.eu/document/download/e3cf4d38-ee41-42f9-8994-f6589ffad458_en?filename=Whitepaper%20AbD_en.pdf
- [30] D. Proença, J. Borbinha, és S. Guerreiro, „eArchiving Capability and Maturity Model”. DLM Forum, 2024. Elérés: 2024. október 29. [Online]. Elérhető: <http://kc.dlmforum.eu/eark-products>
- [31] European Commission. Directorate General for Informatics és CEPS, *Study supporting the final evaluation of the programme on interoperability solutions for European public administrations, businesses and citizens (ISA²): final*

- study*. LU: Publications Office, 2021. Elérés: 2024. október 20. [Online]. Elérhető: <https://data.europa.eu/doi/10.2799/94683>
- [32] S. Segura, L. Gallego, E. Jamin, M. A. Gomez, S. van Hooland, és C. Genin, „ISA² Action 2017.01 Standard-Based Archival Data Management, Exchange and Publication. Study. Final Report”. 2018.
- [33] „DLM Forum”. Elérés: 2024. október 23. [Online]. Elérhető: <https://www.dlmforum.eu/>
- [34] „ArchivesSpace”. Elérés: 2024. november 4. [Online]. Elérhető: <https://archivesspace.org/>
- [35] „AtoM: Open Source Archival Description Software”. Elérés: 2024. november 4. [Online]. Elérhető: <https://www.accesstomemory.org/hu/>
- [36] „Repository Profiles”, Society of American Archivists. Elérés: 2024. november 2. [Online]. Elérhető: <https://www2.archivists.org/groups/collection-management-section/repository-profiles>
- [37] G. Buttars, „Implementing ArchivesSpace | Archives @ University of Edinburgh”. Elérés: 2024. november 3. [Online]. Elérhető: <https://libraryblogs.is.ed.ac.uk/edinburghuniversityarchives/2015/06/22/implementing-archivesspace/>
- [38] „AtoM - Borthwick Institute for Archives, University of York”. Elérés: 2024. november 3. [Online]. Elérhető: <https://www.york.ac.uk/borthwick/projects/atom/>
- [39] „ArchivesSpace”, Github. Elérés: 2024. november 4. [Online]. Elérhető: <https://github.com/archivesspace/archivesspace>
- [40] „Artefactual / AtoM”, Github. Elérés: 2024. november 4. [Online]. Elérhető: <https://github.com/artefactual/atom>
- [41] „ArchivesSpace vs Access to Memory (AtoM) | LibHunt”. Elérés: 2024. november 2. [Online]. Elérhető: <https://selfhosted.libhunt.com/compare-archivesspace-vs-atom>
- [42] „Magyarországi egyházi levéltárak közös segédlete”. Elérés: 2024. november 4. [Online]. Elérhető: <https://arca.melte.hu/>
- [43] „Archivematica: open-source digital preservation system”. Elérés: 2024. november 3. [Online]. Elérhető: <https://www.archivematica.org/en/>
- [44] „RODA Community Edition”, Keep Solutions. Elérés: 2024. november 3. [Online]. Elérhető: <https://www.roda-community.org/>
- [45] „Archivematica Sandbox”. Elérés: 2024. november 3. [Online]. Elérhető: <https://sandbox.archivematica.org/administration/accounts/login/>
- [46] „RODA demo”, RODA Community Edition. Elérés: 2024. november 3. [Online]. Elérhető: <https://demo.roda-community.org/#welcome>
- [47] Zagyva N., „Jegyzetek. Adatbázis-archiválási szakmai gyakorlat”, Google Docs. Elérés: 2024. november 3. [Online]. Elérhető: https://docs.google.com/document/d/17B1hD0oIh3izsb3WMQFkoob7_2wnp-lECufDv7jzU3pw/
- [48] „RODA-In”, Keep Solutions. Elérés: 2024. november 3. [Online]. Elérhető: <https://rodain.roda-community.org/>
- [49] „keeps/roda: RODA - Repository of Authentic Digital Objects”, Github. Elérés: 2024. november 3. [Online]. Elérhető: <https://github.com/keeps/roda>
- [50] „Artefactual/Archivematica”, Github. Elérés: 2024. november 3. [Online]. Elérhető: <https://github.com/artefactual/archivematica>

8. MELLÉKLETEK

8.1 Összegzés

A szakdolgozat bemutatja a Hagyományok Háza Folklór Archívumának előzményeit, intézményi háttérét, gyűjtőkörét, gyűjteményét és vázolja a gyűjteménykezelés jelenlegi folyamatát. Részletezi a jelenlegi gyakorlat hiányosságait és a leginkább megoldandó problémákat.

Ezekre a problémákra keres megoldást a dolgozat, illetve a Hagyományok Háza Adatbázis Szakcsoportja és Gyűjteményi Főosztálya munkatársaiból összeállt projektcsapat, melynek vezetője a szerző.

A dolgozat ismerteti az OAIS referenciamodell, amely fogalmi keretet ad az archívumi rendszerek vizsgálatára, összehasonlítására, és irányt mutat a hosszú távú digitális megőrzés folyamatai és rendszerei megvalósításához. Emellett áttekinti azokat a legfontosabb nemzetközi szabványokat (PREMIS, METS, EAD, ISAD(G), Dublin Core, OAI-PMH), amelyek az archívumi metaadat-kezelés és interoperabilitás alapját képezik.

A szerző kitér a nemzetközi kutatásokra és tapasztalatokra, különösen az európai E-ARK és ISA² programokra. Röviden ismerteti az E-ARK projektek keretében kidolgozott csomagspecifikációkat, a Digitális archiválás referenciaarchitektúráját (eArchiving Reference Architecture 2.0), a tervezett archiválás (Archiving by Design) koncepcióját és az eACMM érettségi modellt. Az ISA² program 2017-es jelentésében szereplő konkrét szoftverajánlások vezettek a következő fejezetben leírt vizsgálatokhoz.

A nyílt forráskódú szoftverek előnyeinek és hátrányainak ismertetését két gyűjteményi adatkezelő rendszer, az ArchivesSpace és az Access to Memory (AtoM), majd két megőrzési rendszer, az Archivematica és a RODA összehasonlítása követi.

A kutatás eredményei megalapozzák a Folklór Archívum digitalizációs projektjének rövid és hosszú távú céljait és terveit.

8.2 Summary

The thesis presents the origins, institutional background, collection scope, and holdings of the Folklore Archive at the Hungarian Heritage House, while outlining the current processes of collection management. It details the shortcomings of current practices and the primary issues that need to be addressed. The project seeks solutions to these issues.

The thesis describes the OAIS reference model, which provides a conceptual framework for examining and comparing archival systems and offers direction in implementing long-term digital preservation processes and systems. It also reviews the essential international standards that form the basis of archival metadata management and interoperability.

The author addresses international research and experiences, particularly the European E-ARK and ISA² programs. It briefly introduces the package specifications developed within the E-ARK projects, the eArchiving Reference Architecture, the concept of Archiving by Design, and the eArchiving Capability Maturity Model. The specific software recommendations in the 2017 ISA² program report lead into the analyses presented in the next chapter.

The discussion of the benefits and drawbacks of open-source software is followed by a comparison of two collection management systems, ArchivesSpace and Access to Memory (AtoM), and two preservation systems, Archivematica and RODA.

The research results form a foundation for the short- and long-term objectives of the Folklore Archive's digitization project.