



Az e-mailek hosszú távú megőrzése

KALCSÓ Gyula – SZÚCS Kata Ágnes

A digitális megőrzés problémái és az e-mail

Az e-mail mint *born digital* objektum hosszú távú (köz)gyűjteményi megőrzése hazánkban új terület. A *Közgyűjteményi Digitalizálási Stratégia* keretében elkészült *Fehér könyv 1.0*¹ még nem tárgyalta, de a megújított második változatába² már bekerül a *born digital* objektumokról szóló fejezetbe. Az e-mail egész életciklusát digitális környezetben tölti; napjaink egyik legáltalánosabban és legszélesebb körben használt kommunikációs formája, a formátum sajátosságai miatt azonban különösen nehéz a kezelése. Az elektronikus levelek és levelezések hosszú távú megőrzési módszertanának részletes kidolgozása sürgető igényként fogalmazódik meg, hiszen már jelenleg is van feldolgozásra váró gyűjteményi tartalom, és a jövőben az ún. LAM-szektor (*Libraries, Archives, Museums* – könyvtárak, levéltárak, múzeumok) intézményei egyre növekvő intenzitással számíthatnak ilyen feladatokra.

A *born digital* témakörében már magyarul és idegen nyelven is elég kiterjedt szakirodalom hozzáférhető,

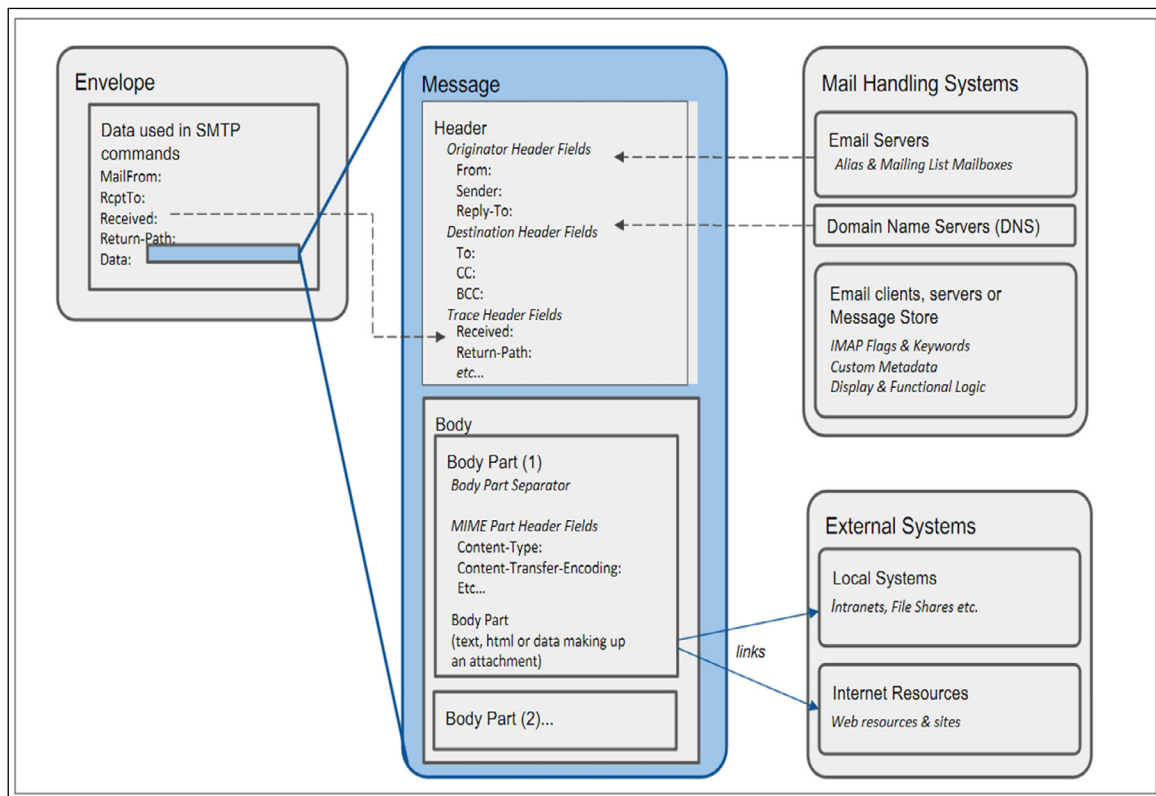
ugyanakkor ezen belül az e-mailek kezelése még nemzetközi szinten is viszonylag kevésbé kidolgozott terület. Ebben éppen az utóbbi néhány évben történt előrelépés. Az *Andrew W. Mellon Foundation* és a *Digital Preservation Coalition* (DPC) 2016-ban hozott létre egy kifejezetten az elektronikus levelek archiválásával foglalkozó munkacsoportot (*Task Force on Technical Approaches for Email Archives*, TFTA EA),³ amely azzal a céllal jött létre, hogy segítse a nemzetközi gyűjteményi közösséget abban, hogy képes legyen kezelni a tudományos és kutatási célokra megőrzendő elektronikus levelezést. A munkacsoport közreadta a *The future of email archives*⁴ [*Az e-mail-archívumok jövője*] című dokumentumot. Ez egy olyan fogalmi és technikai keretet biztosít, amelyben az e-mailek megőrzésére irányuló jelenlegi erőfeszítések nem egymással versengő megoldásokként, hanem egy interoperábilis eszköztár elemeiként működhetnek. Nem sokkal később a DPC is megjelentette *Preserving email* [*Az e-mail megőrzése*] című kiadványát,⁵ amely a 2011-es első

kiadást váltotta fel, és amely a munkacsoport eredményeit összegzi.

Az e-mail mint *born digital* objektum esetében egyszerre merülnek fel a digitális megőrzésre általában jellemző problémák, valamint a kifejezetten az elektronikus levelekre és levelezésekre vonatkozó kérdések. Elsőként érdemes tisztázni a *born digital* fogalmát.

A *born digital* (digitálisan született) meghatározásakor a szakirodalom jó része nem lép túl azon, hogy ezek digitálisan keletkező egységek: a digitális objektumok világában a *born digital* a digitalizálttal állítják szembe. Igaz ugyan, hogy a *born digital* objektumoknak nincs analóg előzményük, viszont ezek sem csak digitális formában létezhetnek (készülhet róluk analóg másolat). Ugyanakkor fontos hangsúlyozni, hogy – miként az analóg objektumokról készült digitális másolatok sem – az analóg változatok soha nem tudják maradéktalanul leképezni a forrásukat. Azaz a *born digital*nak nincs analóg előzménye, de pontos analóg másolata sem lehet, mivel a digitális létmód egyik fontos velejárója, hogy – az *Open Archival Information System* (OAIS: *Nyílt Archívumi Informá-*

ciós Rendszer) referenciamodell szerinti értelemben vett – tartalmi információján túl eleve akár analóg módon reprodukálhatatlan vagy a nem számítógépes közegben értelmezhetetlen metaadatokat is tárol. Egy PDF vagy éppen egy e-mail egyszerű kinyomtatása akár igen nagy számú, a digitális objektumban eredetileg benne lévő, az értelmezhetőség fenntartásához elengedhetetlen metaadat elvesztésével is járhat. Ilyen értelemben a *born digital* objektum – legalábbis az eredeti rendeltetése szerint – a legtöbb esetben *digital exclusive* (azaz kizárólag digitálisan létező) is. A *born digital* kellőképpen árnyalt definiálásában az objektum mint terminus technicus használatának is fontos szerepe van. A *born digital* az esetek nagy részében nem egyetlen fájlból álló egységet jelent. Csakúgy, mint az analóg világ könyvsorozatai, műtárgyegyüttesei, levéltári fondjai, a fájlhalmazok, fájlstruktúrák is igényelhetnek együttes kezelést, sőt: bizonyos esetekben nem is kezelhetők másként. Még ha egy-egy elektronikus levél átkódolt formában akár egyetlen fájlban is tartalmazhatja a csatolt fájlok bitsorozatát, a levélváltások vagy akár teljes levelezés (amennyiben a gyűjteményi szempontok ezt



1. ábra

A szabványos e-mail felépítése. Forrás: PROM 2019 (ld. az 5. jegyzetet), p. 16.

indokolják) együttes megőrzése esetén már egy több fájlból álló egység kezelendő. A *digitális objektum* kifejezés a digitális megőrzésben elterjedt PREMIS metaadatszabvány⁶ szakkifejezése, használata a *Fehér könyv* által is javasolt.

A *born digital* objektumok tehát analóg előzmény nélküli, digitálisan létrejött fájlok vagy strukturált fájlhalmazok, amelyek nem konvertálhatók pontosan (információvesztés nélkül) analóggá. Kezelésük nehézségei egyfelől digitális létmódjukból adódnak, és sok tekintetben nem különböznek a digitális megőrzés általános problémáitól: igen nagy mennyiségben, rendkívüli változatosságban, az értelmezhetőség fenntartását megnehezítő formában (pl. elavult fájlformátumokban) kerülhetnek a gyűjteményekbe. Mielőtt megvizsgálánk, hogy a fenti problémák hogyan vonatkoztathatók az e-mailre mint *born digital* objektumra, valamint hogy ezeken túl milyen speciális nehézségek merülnek fel, célszerű áttekinteni az e-mail technikai hátterét.

Az e-mail felépítését tekintve komplex, és így számos kérdést vet fel. Egyrészt ma már nemcsak egy egyszerű szöveges üzenetről van szó, mint a kommunikációs forma kialakulásakor, hanem többféle tartalom (pl. kép, videó, linkek stb.) és mellékletek kapcsolódhatnak hozzá, továbbá más rendszerekkel (pl. naptár) is össze lehet kötni. Másrészt a szabványos e-mail mint fájl maga is több részből épül fel. A *header* (fejléc) tartalmazza az üzenet szabványban meghatározott metaadatait, a *body* (törzs) pedig az e-mail szöveges részét, illetve a csatolmányokat is. Az e-mailt az *Internet Engineering Task Force* (IETF) 1981 óta szabványosítja. Ezeket a dokumentumokat *Request for Comments*-nek (RFC) hívják [*Felhívás megjegyzésekre*]. A jelenlegi standardot az üzenetformátumra vonatkozóan az *RFC 5322* (IMF – *Internet Message Format*),⁷ az üzenetküldésre vonatkozóan az *RFC 5321* (SMTP – *Simple Mail Transfer Protocol*)⁸ tartalmazza. A fentiek mellett a hozzáférést szabályozó protokollok az *Internet Message Access Protocol* (IMAP)⁹ és a *Post Office Protocol v3* (POP3),¹⁰ továbbá a csatolmányokat és beágyazott, nem szöveges tartalmakat szabályozó *Multipurpose Internet Mail Extensions* (MIME).¹¹

Az *RFC 5322*-ben meghatározottak szerint az e-mailnek alapvetően két szabványos formátuma van, az *.eml* és az *.mbox*. Az *.eml* fájlok minden egyes üzenetet (a fejlécét, a törzsét, beleértve a csatolmányait is) egyetlen fájlként tárolnak. Az *.mbox* formátum hasonló alapon működik, csak egy teljes mappa (nem

feltétlenül egy teljes postafiók) összes üzenetét tárolja el egyetlen fájlban (ún. konténerformátum).

Az e-mailek megőrzésének módjai

Az elektronikus levelezést sokféle szolgáltató teszi lehetővé, de nem létezik univerzális szabványosított eljárás az e-mailek hosszú távú megőrzésére, archiválására. Ezek a kliensalkalmazások különböző formátumokat használnak az egyes e-mailek vagy egész e-mail fiókok exportálására, melyek különböző mértékű beavatkozásokat igényelnek az archiválás során. Az e-mailek kezelésénél többek között figyelni kell arra is, hogy a fájlok és a metaadatok együtt maradjanak. Egy hosszú távú megőrzési rendszer kialakításakor a digitális megőrzés alappilléreire kell támaszkodni, azaz csomagokban szükséges gondolkodni (ld. alább), és egyaránt figyelembe kell venni a formátum egyedi és a digitális objektumokat érintő közös tulajdonságait, ügyelve arra, hogy az eljárás a nemzetközi szabványoknak is megfeleljen.

A DPC által kiadott *Preserving email 2nd edition* (ld. az 5. jegyzetet) című útmutató mellett az e-mailek hosszú távú megőrzésének másik fontos dokumentuma a TFTA EA konzorciumi munkacsoport által kidolgozott jelentés, a *The future of email archives* (ld. a 3–4. jegyzetet). Ebben – a DPC-vel is együttműködve – javaslatot tesznek az e-mail-archiválás standard eljárásaira. A jelentés széles körű célközönségnek készült: a köztintzmények és a GLAM-szektor (*Galleries, Libraries, Archives, Museums* – galériák, könyvtárak, levéltárak, múzeumok) képviselőitől kezdve az informatikusokon és szoftverfejlesztőkön át a kereskedelmi szállítókig és a kutatókig.

Az eddig kialakult nemzetközi gyakorlatok három fő irányt különböztetnek meg az e-mailek hosszú távú megőrzésére:¹²

- A *bitszintű másolat* esetében megmarad az eredeti formátum, amely elsősorban archiválásra alkalmas, ebben a formában nehézkesen lehet szolgáltatni a fájlokat, és a későbbi hozzáférhetőségük is kérdésessé válhat. Szokás a csatolmányokat kibontva, csomagban megőrizni az eredeti fájl mellett.
- A második a *migráció*, amelyen az e-mail-fájlok valamely szabványos és elterjedt formátumra való konvertálását kell érteni. A megőrzendő fájlokat EML-lé vagy MBOX-szá, esetleg tovább, XML-lé – ez utóbbira már van egy kidolgozott ajánlás is, az ún. *Email Account Schema*¹³ – vagy más szabványos formátumra konvertáljuk, majd

repozitóriumba helyezzük. Egészen új elképzelés a TFTA EA munkacsoportjéé, amely egy erre a célra kidolgozott PDF-formátumban képzei el az e-mailek archiválását és hosszú távú megőrzését.¹⁴ A migráció esetében is szokás a csatolmányokat kibontva, csomagban megőrizni az eredeti fájl mellett.

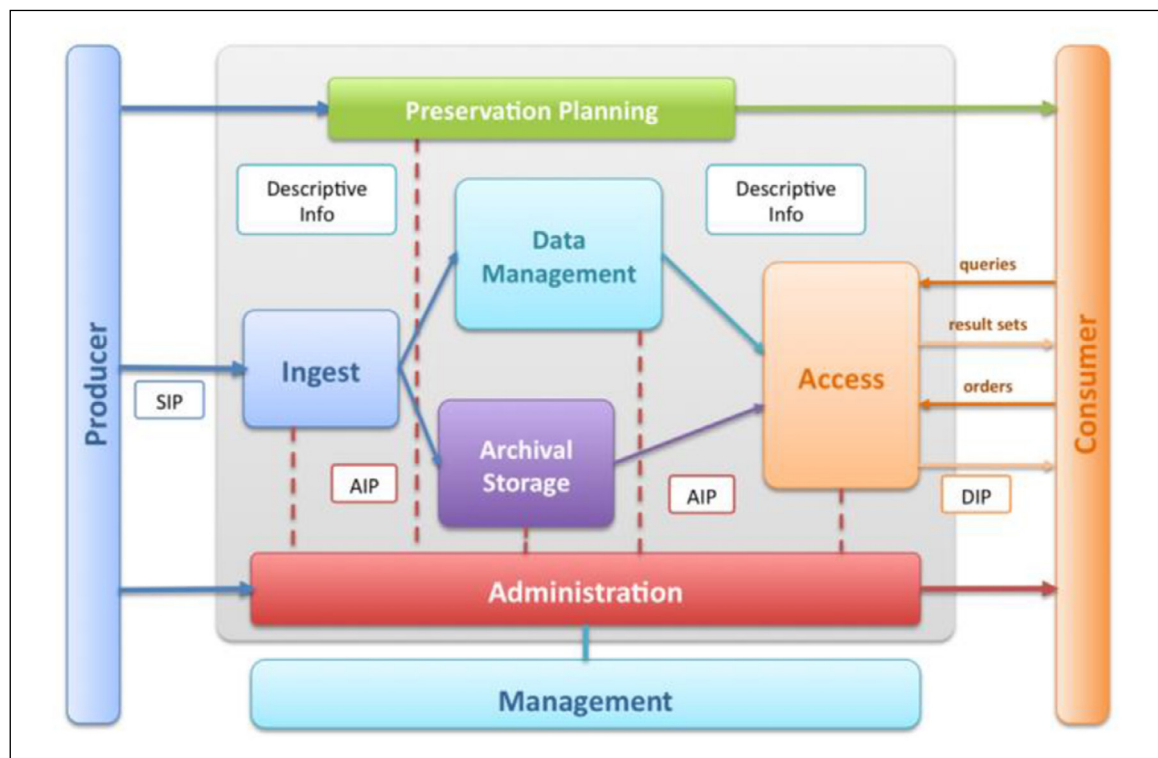
- A harmadik lehetőség az *emulálás*, amely az eredeti szoftveres környezet lehető leghívebb utánzásával vagy annak működő rendszerként történő felépítésével érhető el. Ezt az eljárást azonban általános e-mail-archiválási gyakorlatként nem érdemes használni, mert olyan informatikai igényeket támaszt az intézménnyel szemben, amelyek lassítják a folyamatot. Inkább egy-egy kiemelt projekt esetében lehet érdekes vagy hasznos, illetőleg akkor, ha a használt szoftverkörnyezet olyan kontextuális információt jelent, amely az értelmezhetőséghez fontos vagy elengedhetetlen.

Külön szükséges szólni a célszoftverek használatáról. Az e-mailek kezelésére és archiválására számos kereskedelmi és ingyenes (akár nyílt forráskódú) szoft-

ver áll rendelkezésre. Alkalmazásuk sok szempont mérlegelésétől függhet, a ráfordítható költségkerettől a fenntarthatóság biztosításáig. Nagy hátrányuk, hogy általában viszonylag kis méretű archívumok létrehozására és kezelésére képesek, ezért a nagyobb léptékű archiváláshoz más megoldásra van szükség. Mindenképpen jobb választás az a módszer, amely a lehető legkevésbé teszi függővé a digitális megőrzést speciális szoftverektől és szolgáltatásoktól. Erről lesz szó a következőkben.

Az OAIS és megőrzés csomagban

Az OAIS referenciamodell a digitális archívumok működésének absztrakt leírására létrejött ajánlás.¹⁵ 2012 óta szabványos eljárás,¹⁶ melyet egy NASA-hoz kapcsolódó szervezet (a CCSDS)¹⁷ fejlesztett ki más szervezetekkel együttműködve, és mostanra nemzetközi szinten egyre több intézmény használja (Magyarországon elsőként a Magyar Nemzeti Levéltár).¹⁸ Az OAIS modell sikere egyrészt abban rejlik, hogy megfogalmazza azt a konceptuális keretrendszert, amelynek segítségével a szervezetek elindulhatnak



2. ábra

Az OAIS referenciamodell fő komponensei. Forrás: KÓMÁR – BÁNKI 2019 (ld. az 1. jegyzetet), p. 125.

a hosszú távú megőrzést biztosítani képes archívum létrehozása felé, másrészt pedig ez a magas szintű elméleti modell lefekteti a digitális környezetben keletkező adat és információ hosszú távú megőrzésének minimális követelményeit.

A referenciamodell egyik legfontosabb eleme a csomagokban történő archiválás alapvetővé tétele. Ez azért fontos, mert a digitális objektumok a legtöbb esetben nem egy fájlból állnak, hanem sok összefüggő – és bizonyos esetekben különböző formátumú – fájl alkotja őket, valamint ezzel a módszerrel olyan metaadatok is tárolhatók, amelyek a csomagban található információ hosszú távú, platformfüggetlen értelmezhetőségét biztosítják.

A modell három ágenszt különböztet meg (az információ létrehozója, a megőrző szervezet, valamint a felhasználó), ennek megfelelően három csomagot határoz meg: a létrehozó által benyújtottat, az archiválási célból létrehozottat, valamint a felhasználó számára előállítottat.

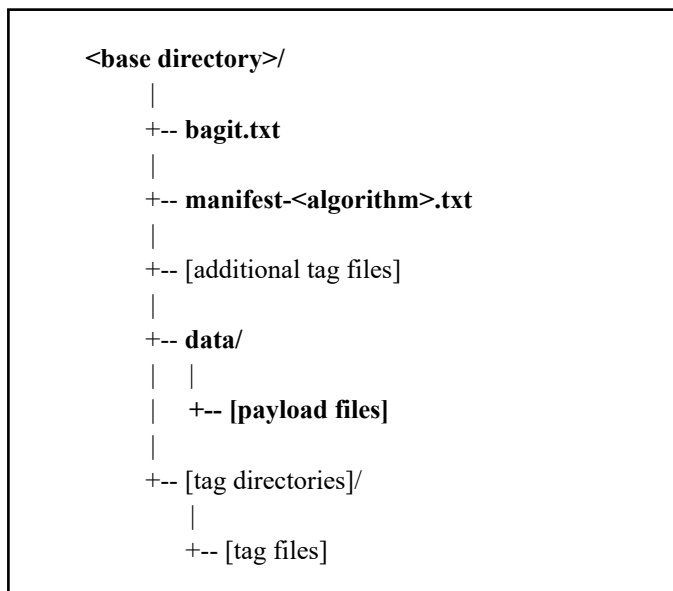
1. *Submission Information Package* (SIP) = átadási információs csomag.
2. *Archival Information Package* (AIP) = megőrzési információs csomag.
3. *Dissemination Information Package* (DIP) = disszeminációs információs csomag.¹⁹

Az OAIS modell középpontjában a menedzsment-feladatok állnak – mint például a megőrzési terv létrehozása –, melyek a különböző adminisztrációs

tevékenységeket, a befogadást és a szolgáltatást is magukban foglalják. A modell emellett előírja azt is, hogy milyen típusú metaadatokot kell az információs csomagoknak tartalmazniuk, ugyanakkor ezek előállításának mikéntjéről nem rendelkezik.

A fent tárgyalt három elektronikuslevél-megőrzési módszer (de az első kettő különösen) megköveteli a megőrzendő digitális objektumok ún. csomagolását, az OAIS-modell alkalmazását. Akár a bitszintű megőrzés, akár valamilyen migrációs megoldás, akár ezek keveréke mellett dönt egy gyűjtemény (de még az emuláció esetében is igaz ez), digitális objektumokat kell megőriznie, amelyek változatos fájlstruktúrában és -típusokként kezelendők. Az OAIS-csomagok létrehozásával garantálható ezek hosszú távú megőrzése, oly módon, hogy a bennük tárolt információ bármikor szoftver- és platformfüggetlenül reprodukálható és értelmezhető marad.

Az OAIS-csomagok felépítésére technikai javaslat a szabványban nincs, hiszen ez egy konceptuális keretrendszer, és (nagyon helyesen) nem kötődik semmilyen fájlrendszerhez vagy -formátumhoz. Emiatt azonban szükség van olyan ajánlásokra is, amelyek a konkrét megvalósítást szabályozzák. A *Library of Congress* által ajánlott *BagIt* egy ilyen csomagolási eljárás, amely megfelel az OAIS-modell elvárásainak is. Tulajdonképpen egy fájlrendezési és -elnevezési konvencióról van szó, amely tetszőleges digitális tartalom tárolását és továbbítását teszi lehetővé.²⁰



3. ábra

A *BagIt* csomagolóeljárás csomagfelépítése (félkövérrel kiemelve a kötelező elemek)

A *BagIt* legfőbb komponense tehát a struktúra, amelynek vannak kötelező tartalmi elemei. A *data* almappában van a *payload*, az ún. hasznos fájlok összessége, azaz maga az adat, amit szeretnénk továbbítani, archiválni. A *bagit.txt* szövegfájl azonosítja, hogy a csomag a *BagIt* konvenciói, illetve annak mely verziója szerint készült. A címkefájlok (*tag*) magát a csomagot írják le, beleértve a fájlok listáját (*manifest-<algorithm>.txt*) és az ellenőrzőösszegeket (*checksum*), melyek a csomagban található fájlok integritását ellenőrzik. A *BagIt* rugalmas rendszere azt is megengedi, hogy a kötelező elemeken kívül további tetszőleges mappákat és fájlokat adjunk hozzá a csomaghoz. Ez tette lehetővé egy kifejezetten e-mailek kezelésére ajánlott csomagtípus kifejlesztését.

A Mailbag

A *Mailbag*²¹ a New York-i állami egyetem (*University at Albany*) M. E. Grenander Különgyűjtemények és Archívumok Tanszékének (*Department of Special Collections and Archives*) kezdeményezése a hosszú távú e-mail-archiválás *BagIt*-en alapuló szabályozására. A projekt 2021 márciusában indult. A specifikáció²² részletesen leírja, hogy a *BagIt*-nek megfelelő elektronikuslevél-archiváló csomagnak mit érdemes tartalmaznia. Az ajánláshoz szoftvert is fejlesztettek,²³ amely alkalmas lesz a *Mailbag*-csomagok létrehozására.

Ajánlásuk szerint a *Mailbag*-csomag a 4. ábra szerinti felépítésű.

Az ún. *payload folder*en (*data*) kívüli fájlok megfelelnek a *BagIt*-specifikációnak, ezen kívül két többletelem van: a *Mailbag*-nek tartalmaznia kell egy *mailbag.csv* fájlt, amelyben *comma separated values* (CSV) formátumban a csomagban tárolt e-mailek bizonyos, az ajánlás által előírt metaadatai szerepelnek; valamint az ugyancsak meghatározott metaadatokat tartalmazó *bag-info.txt* fájlt, amelyet a *BagIt* javasol, a *Mailbag* viszont kötelezővé tesz. A *data folder* legalább egy formátumban tartalmazza a megőrzendő e-mail(ke)t az alábbiak közül: MBOX, PST, MSG, EML, PDF, WARC. Mindegyik formátum a megfelelő nevű mappába kerül (pl. az MBOX az *mbox* nevűbe). Az *attachments* mappa csak akkor kötelező, ha vannak csatolmányok. Amennyiben vannak, az adott e-mailhez hozzárendelődik egy, az üzenet fejlécéből nyert *Message-ID* azonosítóval elnevezett mappa, benne a csatolt fájlokkal.

A *Mailbag* a *payload*-ban található fájlok elnevezésére vonatkozóan is megfogalmaz ajánlásokat. Mind

```
<base directory>/
|
|-- bagit.txt
|
|-- bag-info.txt
|
|-- mailbag.csv
|
|-- manifest-<algorithm>.txt
|
+-- tagmanifest-<algorithm>.txt
|
+-- data/
|
|   |-- mbox/
|   |   |-- [payload files]
|   |-- <optional format1>/
|   |   |-- [payload files]
|   |-- <optional format2>/
|   |   |-- [payload files]
|   |-- ...
|
|   +-- attachments/
|   |   |-- [Mailbag-Message-ID]/
|   |   |-- [payload files]
|   |   |-- [Mailbag-Message-ID]/
|   |   |   |-- [payload files]
|   |   |-- ...
```

4. ábra

A *Mailbag* csomagolójárás csomagfelépítése
(félkövérrel kiemelve a kötelező elemek)

a *BagIt*, mind a *Mailbag* alkalmas az OAIS-csomag szerepére, azaz képezhető belőlük SIP-, AIP- és DIP-csomag, csak a *payload* tartalmát kell megfelelően kialakítani. Az eredeti formátumok SIP-csomagban, a migrált formátumok (pl. a csatolmányok hosszú távú megőrzésre konvertált változatai) az eredetiekkel együtt az AIP-csomagban, a szolgáltatási formátumok pedig a DIP-csomagban szerepelhetnek.

A Petőfi Irodalmi Múzeum tervezett workflow-ja²⁴

A Petőfi Irodalmi Múzeum Digitális Bölcsészeti Központja kidolgozott egy eljárásrendet a múzeumi gyűjteményi e-mailek archiválására. A folyamat legfontosabb célkitűzése a múzeumi e-mailek szakszerű

és szabványos archiválása. Mivel a fenti ajánlásokon alapul, ezért egyaránt alkalmas az adattári célú dokumentációs gyűjteményi elemek, de akár írói elektronikus levelezések kezelésére is.

Az elképzelés lényege, hogy az OAIS-nak megfelelő csomagok beilleszthetők legyenek a meglévő integrált múzeumi gyűjteménykezelő rendszerbe. Ez azonban csak egy olyan szoftverkomponens közbeiktatásával volt lehetséges, amely képes a csomagok automatikus előállítására és kezelésére.

A pilotprojekt során egy teljes postafiók tartalmának gyűjteményi archiválása volt a feladat. A folyamat első lépése a gyűjteményezési folyamat részeként elvégzett válogatás, valamint a szükséges metaadatfájlok automatikus előállítása. A válogatás során manuálisan eltávolításra kerülnek a feldolgozandó fiókból a felesleges mappák (spam, törölt üzenetek), valamint az esetleg spamszűrőn átment üzenetek. A pilotprojektben a válogatás után exportált PST-fájlból a *Bitcurator*-be épített konverziós eszköz segítségével (*readpst*) MBOX-fájl készül, amelyből kibonthatók az EML-ek. A forrásfájlokból kibontott és szükség esetén EML-lé konvertált üzenetek megfelelő mappastruktúrába kerülnek egy *Python-script*²⁵ segítségével, ez kiolvassa az e-mailek *header*jéből a metaadataikat is, valamint beírja azokat a megfelelő metaadatfájlba. A válogatás részeként vírus- és *malware*-ellenőrzés is történik (a fertőzött állományok eltávolításra kerülnek), az erről készült reportfájl bekerül a SIP-csomagba.²⁶ Ezután a *RODA-In*²⁷ nevű eszközzel elkészül a SIP-csomag.

Jegyzetek és irodalmi hivatkozások

1. KÓMÁR Éva – BÁNKI Zsolt, szerk. Fehér könyv: Módszertani útmutató a közgyűjteményi kulturális örökség digitalizálásához és közzétételéhez [elektronikus dok.]. Budapest: Emberi Erőforrások Minisztériuma, 2019. 258 p. Hozzáférhető: http://www.oszk.hu/sites/default/files/Feher_Konyv.pdf [Megtekintve: 2021.12.08.] További hozzáférés: https://mandadb.hu/dokumentum/1096777/feher_konyv.pdf [Megtekintve: 2021.12.08.] További hozzáférés: http://mnl.gov.hu/sites/default/files/feher_konyv_0.pdf [Megtekintve: 2021.12.08.]
2. Megjelenés alatt.
3. A munkacsoport működéséről jó összefoglalót ad: PROM, Chris – MURRAY, Kate. Task Force on Technical Approaches to Email Archives: DPC Briefing Day: July 6, 2017, The National Archives, Kew [elektronikus dok.]. = A Digital Preservation Coalition honlapja. Dokumentumtár aloldal. Digital Preservation Coalition, cop. 2022. Hozzáférhető: <https://www.dpconline.org/docs/miscellaneous/events/2017->

A SIP-csomag előállításának részeként keletkeznek a csomag *checksum*jai, valamint bekerül a csomagba egy METS-²⁸ és egy PREMIS-fájl is a strukturális és adminisztratív metaadatok tárolására. A létrejövő csomag egy tömörített állomány.

Ezen a ponton szükség van egy olyan szoftver alkalmazására, amely be tudja fogadni a SIP-csomagot, előállítja belőle a megfelelő AIP-csomago(ka)t, valamint minden elvégzett műveletet naplóz (beírja a PREMIS-fájlokba). Minderre több lehetőség is van, melyek közül a RODA²⁹ nevű ingyenes és nyílt forráskódú repozitóriumszoftver talán a legjobb megoldás, mivel képes a fenti követelmények teljesítésére, valamint kiváló REST API-val rendelkezik, ami alkalmassá teszi arra, hogy együttműködjön más rendszerekkel. A RODA-val előállított AIP-csomagokkal további műveletek végezhetőek (a hosszú távú megőrzési formátumok bekerülhetnek a csomagokba stb.), ezekről PREMIS-bejegyzések készülnek. Amennyiben szükséges, a szoftver létre tudja hozni a megfelelő tartalmú szolgáltatási csomagot (DIP) is. A tervek szerint a PIM rendszerei (*Médiaweb*, *Huntéka*) elérik a RODA API-ját, és a segítségével előállítják a megfelelő tartalmakat. Az OAIS-csomagok tömörített állományként kerülnek a múzeumi rendszerekbe, azokból szükség esetén a RODA állítja elő a szolgáltatandó tartalmat. Megjegyzendő ugyanakkor, hogy az e-mailek szolgáltatása meglehetősen problematikus terület, számos tisztázásra váró jogi és egyéb kérdéssel, amelyek tárgyalása már túlmutat jelen cikk keretein.

[events/1697-2017-july-email-prom-and-murray-session-1](https://www.loc.gov/standards/premis/) [Megtekintve: 2021.12.08.]

4. The future of email archives: A report from the Task Force on Technical Approaches for Email Archives [elektronikus dok.]. Washington DC: Council on Library and Information Resources, 2018. 120 p. ISBN 978 1 932326 59 8. Hozzáférhető: <https://www.clir.org/wp-content/uploads/sites/6/2018/08/CLIR-pub175.pdf> [Megtekintve: 2021.12.08.]
5. PROM, Christopher J. Preserving email [elektronikus dok.]. 2. kiad. Digital Preservation Coalition, 2019. 51 p. ISSN 2048-7916. Hozzáférhető: <http://doi.org/10.7207/twr19-01> [Megtekintve: 2021.12.08.]
6. PREMIS: Preservation metadata [online]. = A Library of Congress honlapja. Szabványok aloldal. Hozzáférhető: <https://www.loc.gov/standards/premis/> [Megtekintve: 2021.12.08.]

7. IMF: Internet Message Format: RFC 5322 [online]. IETF Trust, 2008. Hozzáférhető: <https://datatracker.ietf.org/doc/html/rfc5322> [Megtekintve: 2021.12.08.]
8. SMTP: Simple Mail Transfer Protocol: RFC 5321 [online]. IETF Trust, 2008. Hozzáférhető: <https://datatracker.ietf.org/doc/html/rfc5321> [Megtekintve: 2021.12.08.]
9. IMAP: Internet Message Access Protocol – Version 4rev1: RFC 3501 [online]. IETF Trust, 2003. Hozzáférhető: <https://datatracker.ietf.org/doc/html/rfc3501> [Megtekintve: 2021.12.08.]
10. POP3: Post Office Protocol – Version 3: RFC 1939 [online]. IETF Trust, 1996. Hozzáférhető: <https://datatracker.ietf.org/doc/html/rfc1939> [Megtekintve: 2021.12.08.]
11. A sztenderd MIME-t a következő RFC-kben határozták meg: RFC 2045, RFC 2046, RFC 2047, RFC 4288, RFC 4289 és RFC 2049.
12. PROM, i. m., p. 15.
13. FERRANTE, Riccardo – SCHMITZ FUHRIG, Lynda. Digital preservation: Using the email account XML schema [elektronikus dok.]. == Archiving 2009: May 4–7, 2009, Arlington, Va.: Preservation strategies and imaging technologies for cultural heritage institutions and memory organizations: Final program and proceedings. Society for Imaging Science and Technology, 2009. 8 p. Hozzáférhető: https://siarchives.si.edu/sites/default/files/pdfs/Archiving2009_EmailAccountSchema_0.pdf [Megtekintve: 2021.12.08.]
14. A specification for using PDF to package and represent email [online]. University of Illinois at Urbana-Champaign, cop. 2021. 34 p. Hozzáférhető: <https://zenodo.org/record/5747194> [Megtekintve: 2022.09.01.]
15. Recommendation for space data system practices: Reference model for an Open Archival Information System (OAIS): Magenta book, June 2012 [elektronikus dok.]. Washington, DC: Consultative Committee for Space Data Systems, 2012. 135 p. (Recommended practice 2., CCSDS 650.0-M-2). Hozzáférhető: <https://public.ccsds.org/pubs/650x0m2.pdf> [Megtekintve: 2021.12.08.]
16. ISO 14721:2003: Space data and information transfer systems: Open archival information system: Reference model [online]. == Az ISO honlapja. Standards catalogue oldal. Geneva: ISO, [s.a.]. Hozzáférhető: <https://www.iso.org/standard/24683.html> [Megtekintve: 2021.12.08.]; illetve ISO 14721:2012: Space data and information transfer systems: Open archival information system (OAIS): Reference model [online]. == Az ISO honlapja. Standards catalogue oldal. Geneva: ISO, [s.a.]. Hozzáférhető: <https://www.iso.org/standard/57284.html> [Megtekintve: 2021.12.08.]
17. The Consultative Committee for Space Data Systems (CCSDS) honlapja [online]. [S.I.]: CCSDS/ASRC Federal Agile Decision Sciences, cop. 2022. Hozzáférhető: <https://public.ccsds.org/default.aspx> [Megtekintve: 2021.12.08.]
18. OAIS – Open Archival Information System [online]. A Magyar Nemzeti Levéltár honlapja. Szabványok oldal. Budapest: MNL, [s.a.]. Hozzáférhető: <https://mnl.gov.hu/mnl/szkk/szabvanyok> [Megtekintve: 2021.12.08.]
19. Az MSZ ISO 14721 megjelenés alatt álló magyarítása szerinti kifejezések.
20. The BagIt File Packaging Format (V1.0): RFC 8493 [online]. IETF Trust, 2018. Hozzáférhető: <https://datatracker.ietf.org/doc/html/rfc8493> [Megtekintve: 2021.12.08.]
21. A Mailbag projekt weboldala. M. E. Grenander Department of Special Collections & Archives, University at Albany, [s.a.]. Hozzáférhető: <https://archives.albany.edu/mailbag/> [Megtekintve: 2021.12.08.]
22. A Mailbag-specifikáció [online]. == A Mailbag projekt weboldala. M. E. Grenander Department of Special Collections & Archives, University at Albany, [s.a.]. Hozzáférhető: <https://archives.albany.edu/mailbag/spec/> [Megtekintve: 2021.12.08.]
23. A Mailbag projekt GitHub-repozitóriuma [online]. == A GitHub honlapja. M. E. Grenander Department of Special Collections & Archives, University at Albany oldal. GitHub, cop. 2022. Hozzáférhető: <https://github.com/UAlbanyArchives/mailbag> [Megtekintve: 2021.12.08.]
24. Az előadás időpontjában (2021. november 24-én) tervezett workflow. Időközben szervezeti változások miatt a megvalósulás OSZK–PIM együttműködésben történhet meg, és részleteiben az előadottaktól eltérést mutathat.
25. A bagit-python nevű script implementációja [online]. == A GitHub honlapja. Library of Congress oldal. GitHub, cop. 2022. Hozzáférhető: <https://github.com/LibraryOfCongress/bagit-python> [Megtekintve: 2021.12.08.]
26. Az előfeldolgozás részben automatizálható a BitCurator nevű Linux-disztribúció használatával. Ld. BitCurator [számítógépes szoftver]. [S.n.], [s.a.]. Hozzáférhető: <https://bitcurator.net/> [Megtekintve: 2021.12.08.]. Ennek előre telepített eszközei számos művelet elvégzését teszik lehetővé: PST konvertálását MBOX-szá, EML-ek kibontását MBOX-ból, vírus- és malware-ellenőrzést reportfájlok előállításával stb.
27. A RODA-In weboldala. Keep Solutions, [s.a.]. Hozzáférhető: <https://rodain.roda-community.org/> [Megtekintve: 2021.12.08.]
28. METS: Metadata Encoding and Transmission Standard [online]. == A Library of Congress honlapja. Standards oldal. Washington: Library of Congress, 2022. Hozzáférhető: <https://www.loc.gov/standards/mets/> [Megtekintve: 2022.09.01.]
29. A RODA weboldala. Keep Solutions, [s.a.]. Hozzáférhető: <https://www.roda-community.org/> [Megtekintve: 2021.12.08.]

(Beérkezett: 2022. február 17.)