

Failivormingute nõuded pikaajaliseks säilitamiseks

Eesti rahvusraamatukogu
Digitaalarhiivi osakond

Kadri Varipuu
Jaan Krupp
Eva-Maria Artus

Sisukord

Failivormingute nõuded pikaajaliseks säilitamiseks.....	1
Sissejuhatus.....	3
Eelistatud failivormingud.....	3
Nõuded säilitatavatele failidele.....	3
Üldised nõuded failidele.....	3
Tekstiline materjal.....	4
Üldised nõuded PDFFile:	4
PDF/A.....	4
EPUB ja MOBI	5
XML.....	5
TXT	6
Pildiline materjal.....	6
TIFF	6
JPEG 2000	7
PDF/A pildifailidele	8
Heliline materjal	8
Videomaterjal.....	10
Eelistatud vormingud	10
Veebisaidid	11
Muud failid ja failivormingud	11

Sissejuhatus

Käesoleva dokumendi eesmärk on kinnitada Rahvusraamatukogu digitaalarhiivis pikaajaliseks säilitamiseks sobilikud failivormingud. Kuna Säilituseksemplari seaduse jõustumisega loovutatakse lisaks trükistele ka nende algmaterjal ehk väljaande avaldamiseks loodud digitaalne andmekogum, siis on käesoleva dokumendi üheks eesmärgiks anda algmaterjali loovutamiseks täpsemad juhised. Peale loovutatava materjali leidub digitaalarhiivis ka skaneeritud teoseid, helifaile, videofaile ja veebiarhiivi faile.

Dokumendis esitatud kriteeriumite loomisel on võetud aluseks nii Rahvusraamatukogu kui ka rahvusvaheliste raamatukogude ja arhiivide senine praktika. Samuti on vaadeldud, millised tendentsid pikaajalises digitaalses säilitamises on lähitulevikus ootuspärased. Failivormingute valimisel on arvestatud sellega, millist tüüpi sisu Rahvusraamatukogu digitaalarhiivis leidub.

Nõuded vaadatakse igal aastal uuesti üle ning kooskõlastatakse erinevate Rahvusraamatukogu osakondadega, mis tegelevad digitaalse materjaliga.

Eelistatud failivormingud

Käesolevas dokumendis loetletud failivormingud tagavad arhiveeritavate failide ja nendes oleva sisu kättesaadavuse ning kasutatavuse tulevikus. Failivormingud, mis ei ole laialdaselt toetatud ega levinud, ei pruugi tulevikus enam avaneda ning nendes olev sisu võib jääda kas osaliselt või täielikult kättesaamatuks.

Eelistatud on avatud vormingud, mille kirjeldused on kõigile kättesaadavad ja vabalt tõlgitavad. Avatud failivorming tähendab, et faili struktuur on vastavuses mitmete ettevõtete ja avalike organisatsioonide koostöös valminud standardiga, spetsifikatsioon on kõigile soovijaile kättesaadav. Avatud failivorming peab olema sõltumatu ühest kindlast rakendusprogrammist, tootjast või operatsioonisüsteemist. Vorming peab olema kirjeldatud (ja soovitatavalt standarditud) viisil, mis võimaldab selle täielikku, litsentsivaba kitsendusteta rakendamist kõikidel soovijatel¹.

Nõuded säilitatavatele failidele

Järgnevalt toodud nõuete täitmine on eelduseks, et failidega oleks võimalik teostada säilitustegevusi. Tulevikus ei pruugi tänapäevased failivormingud enam avatavad ja kasutatavad olla ning seega peab olema võimalik kadudeta või minimaalsete kadudega faile konverteerida ühest vormingust teise.

Üldised nõuded failidele

- Faili sisu peab vastama võimalikult üks ühele trükisele, sh sisaldama kaasi, kõiki lehekülgi ja lisasid.
- Fail peab sisaldama teose ametlikult väljaantud lõplikku versiooni.
- Faili sisu peab olema võimalikult kõrges kvaliteedis.
- Fail võib olla kadudega tihendatud vormingus (nt MP3, AAC) vaid juhul, kui teost kõrgemas kvaliteedis ei eksisteeri.
- Fail ei tohi olla kokkupakitud.

¹ https://et.wikipedia.org/wiki/Avatud_dokumendivormingud (23.03.2017)

- Fail ei tohi olla kaitstud parooliga.
- Fail ei tohi olla krüpteeritud.
- Failil ei tohi olla kasutusõigusi piiravat tehnoloogiat (nt DRM)

Tekstiline materjal

Üldised nõuded PDFile:

- PDFis sisalduv tekst peab olema masinloetav (st mitte pilt). Juhul kui tekst ei ole masinloetav, on sunnitud RR failile tegema tekstituvastuse ning säilitama faili sellisena.
- PDF failil ei tohi olla peal kirjutuskaitse.
- PDF faili lehekülgede lahtilõikamine peab olema lubatud.
- PDF faili printimine peab olema lubatud.
- PDF faili sisu kopeerimine peab olema lubatud.
- PDF faili sisu muutmine peab olema lubatud.
- PDFi sisse ei tohi olla lisatud videoid, muusikat, programme ja skripte, sest RR ei suuda tagada nende töökindlust tulevikus.
- Soovitav on PDFile panna fondid kaasa, sest vastasel juhul ei saa RR tagada õige kirjastiili digitaalset säilitamist.

Rahvusraamatukogus säilitatakse praegu trükiseid ja pisitrükiseid PDF 1.4 vormingus.

PDF/A

PDF/A on arhiveerimiseks ja pikaajaliseks säilitamiseks väljatöötatud failivorming. PDF/A erineb tavalisest PDFist selle poolest, et keelatud on pikaajaliseks säilitamiseks sobimatud tunnused, nt fontide (kirjastiilide) linkimine ja krüpteerimine. PDF/A faili vaaturi nõuetes on ka värvide haldamise juhised, põimitud (*embedded*) fontide tugi ja kasutajaliides põimitud märgete (*embedded annotation*) lugemiseks.

PDF/A kindlustab, et elektroonilisi dokumente oleks võimalik taasesitada erinevate tarkvaradega samasuguses vormis, nagu nad algselt loodi. Info, mida on vaja faili taasesitamiseks, on failiga kaasa pandud. See tähendab, et PDF/A dokumendid on 100% iseseisvad ning see kehtib kogu sisu kohta (tekst, rastergraafika, vektorgraafika), fondid ja värvid. PDF/A vorming ei ole sõltuv välistest allikatest (nt fondiprogrammid ja andmevood/*datastream*), kuid võib sisaldada märkmeid (nt lingid) välistele dokumentidele. PDF/A sisaldab ka mitmesuguseid metaandmeid (nt autoriõigused).

Nimi	PDF (<i>Portable Document Format</i>), alamtüüp PDF/A-1b
Failinime laiend	.pdf
Standard	<i>ISO 19005-1:2005 Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4, Level B Conformance (PDF/A-1b)</i> https://www.iso.org/standard/38920.html

Pikaajaliseks digitaalseks säilitamiseks soovitame kasutada vormingut **PDF/A-1b²**.

Nõuetele vastavuse tasandi B nõudmised on minimaalselt vajalikud tagamaks, et nõuetele vastavuses oleva faili esitatav välimus on pikaajaliselt säilitatav. Siiski ei pruugi nõuetele vastavuse tasandi B failides olla piisavalt palju sisemist informatsiooni, mis lubaks samuti kui nõuetele vastavuse tasandil A säilitada dokumendi loomulikus järjestuses lugemiseks vajalikku loogilist struktuuri ja sisu tekstivooge³.

Praegu loovutatakse Rahvusraamatukogule enamik materjale PDF-vormingus (mitte PDF/A). Hiljemalt uude digitaalarhiivi andmete migreerimisel tuleb säilitatavad failid üle viia PDF/A-1b vormingusse.

EPUB ja MOBI

Rahvusraamatukogu võtab vastu ka EPUB- ja MOBI-vormingus väljaandeid, kuid nende pikaajaline säilimine ei ole hetkel tagatud. EPUB-vormingul on suurim potentsiaal saada avatud ja toetatud e-teaviku standardiks, kuid senised versioonid ei ühildu tagasiulatavalt⁴, mis tähendab, et uuemaid EPUBi versioone ei saa konverteerida vanemateks versioonideks.

Hiljemalt uude digitaalarhiivi andmete migreerimisel tuleb EPUB ja MOBI failid üle viia säilitusvormingusse. Täpsed plaanid selleks selgitab välja uue digitaalarhiivi eelanalüüs.

XML

XML on W3C (*World Wide Web Consortium*) väljatöötatud ja soovitatud standardne üldotstarbeline märgistuskeel, mille eesmärgiks on struktureeritud info jagamine infosüsteemide vahel. XML on inim- ja masinloetav keel. Dokumendi struktuuri puudutavad andmed diferentseeritakse selgelt, mis tuleb kasuks dokumentide otsingul, automaattöötusel ja andmehaldusel.

XML fail on vaja esitada:

- reeglipärasel kujul (*well-formed*) vastavalt XML 1.0 spetsifikatsioonile;
- seotult ning valideeruvalt sobiva XML skeemiga (*XML Schema*);
- seotult XML laadileh(ted)ega, mis on valiidne W3C soovitusel mõttes (*W3C Extensible Stylesheet Language (XSL) Version 1.1*)⁵

RR digitaalarhiivis on XML vormingus metaandmete failid, näiteks kirjeldatakse XML failides väljaannete struktuuri (METS.xml) ja tekstivastust.

Nimi	XML (<i>Extensible Markup Language</i>), 1.0
Failinime laiend	.xml

² Sama vormingut soovitab ka [Arhiivieskirja määruse](https://www.riigiteataja.ee/aktiis/1130/1201/5016/VV_2m_lisa.pdf#) Lisa 1 https://www.riigiteataja.ee/aktiis/1130/1201/5016/VV_2m_lisa.pdf# (06.03.2017)

³ <http://www.digar.ee/arhiiv/nlib-digar:213601> (09.03.2017)

⁴ http://wiki.dpconline.org/images/0/04/EBook_Summary_Assessment_v1.0.pdf (09.03.2017)

⁵ https://www.riigiteataja.ee/aktiis/1130/1201/5016/VV_2m_lisa.pdf (11.04.2017)

Standard	<i>Extensible Markup Language (XML) 1.0 (Fifth Edition)</i> <i>W3C Recommendation 26 November 2008</i> https://www.w3.org/TR/xml/
-----------------	--

TXT

Rahvusraamatukogu digitaalarhiiv säilitab ka lihtteksti faile (.txt), mis esitatakse UTF-8 (Unicode) ISO 10646 märgikoodis. Tekstifail (inglise keeles tuntud kui "*plain text*" või lihtsalt "*text file*") on lihtsaim elektroonilise tekstidokumendi vorming, sisaldades vaid tähti, kirjavahemärke, tühikuid ja reavahetusi.

Hetkel kasutatakse arhiivis seda vormingut helisalvestistel olevate lugude loetlemiseks, et neid oleks võimalik pealkirja alusel otsida.

Nimi	Text file, lihtteksti fail
Failinime laiend	.txt
Standard	puudub

Pildiline materjal

TIFF

TIFF on rasterkujutiste tihendamata või kadudeta tihendamise failivorming. Vormingut kasutatakse laialdaselt rasterpiltide (eriti fotode) salvestamiseks ja edastamiseks tarbograafika, kirjastamise ja polügraafia ettevõtetes, sest toetab CMYK-värvimudelit (suurim värvisügavus kuni 32 bitti värvikomponendi kohta). Samuti on võimalik töödelda RGB-signaale (vastavalt 24 bitti).

TIFFi rakendusvaldkonda kuulub samuti dokumentide (sealhulgas nt ühevärviliste tehniliste jooniste) arhiveerimine, sest nii säilivad nad edaspidiseks kasutamiseks ja töötlemiseks algsel kujul. Kuna TIFF-failid on pakkimata või ainult vähesel määral (nt neljakordselt) kadudeta tihendatud, siis on nad palju andmemahukamad kui näiteks vastavad JPEG-failid.

Nimi	TIFF (<i>Tagged Image File Format</i>), Revision 6.0
Failinime laiend	.tif, .tiff
Standard	ISO 12639:2004 (Teine väljaanne) standard TIFF/IT https://www.iso.org/standard/34342.html Lisandused: ISO 12639:2004 / Amd. 1:2007

TIFF sobib säilitamiseks, kuna:

- kõige laialdasemalt toetatud vorming trükitööstuses
- ühildub kõikide arvutiplatvormidega. Sobib erinevatele aplikatsioonidele, on sõltumatu arhitektuurist, operatsioonisüsteemist, graafikast ja riistvarast.
- väga lai värvipalett (full RGB)
- parem pildi skaleerimine⁶ kui JPG-I või GIF-il
- sobib hästi pilditöötamiseks
- töötab nii PostScript kui mitte-PostScript printeritega
- parim vorming arhiveerimiseks⁷

TIFF failidena säilitatakse praegu Rahvusraamatukogus digiteeritud materjale, fotosid ja kaarte.

Pikaajaliseks säilitamiseks soovitame kasutada TIFF_UNC (*uncompressed bitmap*), TIFF 6.0, lossless⁸.

Nõuded TIFFile:

- TIFF v6.0 baseline
- TIFF_UNC versioon 6.0 RGB värvilistele piltidele
- TIFF_UNC versioon 6.0 baseline grayscale mustvalgetele piltidele
- LZW tihendus
- Üks pilt faili kohta ehk et fail ei tohi olla mitmeleheküljeline (multipage).
- Fail ei tohiks sisaldada muid kaasapandud (embedded) vorminguid (nt ZIP, JP2 vms).
- Kui originaalteos on värvipilt (foto, kaart, värviline tekst, digitaalselt sündinud RGB formaadis teos), tuleb seda ka säilitada Full RGB TIFF failina.
- Kui skaneeritakse mustvalgeid dokumente, nt mustvalged kaardid, trükitud dokumendid (mida on raske lugeda ja kus võib olla pliiatsimärkmeid), käsikirjalised dokumendid või digitaalselt sündinud dokumendid, mis on loodud halltoonides, tuleb need salvestada TIFF grayscale failina.

NB! Täpsemad tehnilised parameetrid (lahutusvõime, värvisügavus) TIFF failidele on Eesti Rahvusraamatukogu digiteerimise põhimõtetes⁹.

JPEG 2000

Aktsepteeritavaks failivorminguks pildimaterjali puhul on JPEG 2000, kuna teatud juhtudel on see vorming sobivam kui TIFF. Kui tegemist on värvilise digiteeritud materjaliga, siis võivad failid TIFF vormingu kasutamisel mahult liiga suureks muutuda ja nende töötlemine muutub seetõttu aeglaseks. JPEG 2000 võimaldab faile kadudeta tihendada sealjuures säilitades kvaliteeti ning kahandades failisuurust.

Kuigi vormingul on oma erinevate omaduste ja funktsionaalsuste tõttu potentsiaali olla sobiv vorming pikaajaliseks säilitamiseks, ei ole vorming väljaspool arhiive laialdaselt levinud ega ka toetatud. Lisaks võib failide loomisel või tihendamisel vigu tekkida ning on piiratud hulk võimalusi, kuidas failide

⁶ Fotosid on tarvis skaleerida siis, kui soovitakse neid esitada arvutiekraanil kindlates tingimustes. Vt ka <http://www.eki.ee/dict/ametnik/index.cgi?Q=skaleerima> (24.03.2017)

⁷ <http://ec.europa.eu/ipg/standards/image/tiff/> (23.03.2017)

⁸ <http://www.digitalpreservation.gov/formats/fdd/fdd000023.shtml> (09.03.2017)

⁹ https://www.nlib.ee/public/documents/juhendid-eeskirjad-hinnad-jne/RR_dig_pohimotted_2015.doc (23.03.2017)

valiidsust kontrollida. Võrreldes TIFFiga on tegemist ka tehnoloogiliselt keerukama vorminguga, mis tähendab, et selliste failide töötlemiseks ja lugemiseks on vaja rohkem riist- ja tarkvaralist jõudlust¹⁰.

Nimi	JPEG 2000
Failinime laiend	.jp2
Standard	ISO/IEC 15444-1:2016. <i>Information technology -- JPEG 2000 image coding system: Core coding system</i> : https://www.iso.org/standard/70018.html

Nõuded JPEG 2000-le:

- JPEG 2000 sisse ei tohi olla lisatud videoid, muusikat, programme ja skripte, sest RR ei suuda tagada nende töökindlust tulevikus.

PDF/A pildifailidele

RR ei soovita esmase valikuna pildifaile säilitada PDF-vormingus. Kui algmaterjal on PDF, siis tuleks fail viia säilitamiseks üle PDF/A-1b vormingusse.

Heliline materjal

Helilise materjali säilitamiseks sobib kõige paremini WAVE. Selle vormingu Broadcast versioon on eelistatud, kuna võimaldab metaandmeid lisada faili külge. WAVE on lihtsa struktuuriga vorming, mis edastab heli kompressseerimata ning on väga laialdaselt kasutusel. Broadcast WAVE faile saavad kasutada samad aplikatsioonid, mis võimaldavad ka tavalisi WAVE faile kasutada. Uusim ja eelistatuim versioon BWF-st on versioon 2.0.

Aktsepteeritavateks vorminguteks on valitud MP3 (MPEG-1 Layer 3 ja MPEG-2 Layer 3) ja AAC (*Advanced Audio Coding*). MP3 on väga laialdselt levinud vorming, mis võimaldab heli edastada kokkupakitult ja kadudega sealjuures säilitades heli kvaliteeti. Parima heli saavutamiseks on soovitatav kasutada avatud lähtekoodiga LAME kodeerijat. AAC on Apple'i loodud vorming, mis on patenditud ja vähemlevinud kui MP3, kuid tööme selle vormingute dokumendis välja, sest iTunes Music Store on üks levinumaid viise, kuidas tänapäeval digitaalset muusikat levitatakse ja ostetakse.

NB! Säilitame ka helifailide juurde kuuluvaid illustratsioone (albumi esikaas, heliplaadi kujunduse failid) eelistatult TIFF või PDF vormingus (täpsemad kirjeldused vormingute kohta pildilise materjali peatükis).

Eelistatud failivormingud ja minimaalsed tehnilised parameetrid:

Vorming ja laiend	Kooder	Diskreteerimis-aste	Diskreteerimis-sagedus	Bitsügavus	Standard
-------------------	--------	---------------------	------------------------	------------	----------

¹⁰ <https://blogs.loc.gov/thesignal/2013/01/is-jpeg-2000-a-preservation-risk/> (24.03.2017)

Broadcast Wave Format (.wav)	Linear Pulse Code Modulated Audio (LPCM)	1,411.2 kbit/s	44.1 kHz	16 bit	<i>Specification of the Broadcast Wave Format (BWF) - Version 2.0:</i> https://tech.ebu.ch/docs/tech/tech3285.pdf
Waveform Audio (.wav)	Linear Pulse Code Modulated Audio (LPCM)	1,411.2 kbit/s	44.1 kHz	16 bit	<i>Multiple channel audio data and WAVE files:</i> https://msdn.microsoft.com/en-us/library/windows/hardware/dn653308(v=vs.85).aspx

Aktsepteeritavad vormingud ja minimaalsed parameetrid

Vorming ja laiend	Kooder	Diskreterimisaste	Diskreterimisagedus	Standard
MPEG-1 Layer 3, MPEG-2 Layer -3 (.mp3)	LAME	320 kbit/s	44.1 kHz	<i>ISO/IEC-11172-3: 1993, Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 3: Audio:</i> http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22412 <i>ISO/IEC 13818-3:1998 Information technology – Generic coding of moving pictures and associate audio information – Part 3: Audio:</i> http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=26797
MPEG-4 AAC Advanced Audio Coding (.m4a)		256 kbit/s	44.1 kHz	<i>ISO/IEC 14496-3:2009, Information technology - Coding of audio-visual objects - Part 3: Audio:</i> http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=53943

Videomaterjal

Motion JPEG 2000 kasutatakse paljude arhiivide poolt (sageli koos MXF wrapper'iga), kuna tegemist on avatud standardiga ja võimaldab kadudeta kokku pakkida. MFX wrapper on avatud lähtekoodiga ja laialdaselt kasutusel, kuna võimaldab metaandmeid lisada videofaili külge. Olenevalt sellest, kuidas on faili kodeeritud, ei pruugi MXF failid nii öelda traditsioonilises tähenduses mängida, sest nad on mõeldud professionaalseks kasutamiseks, mitte arvutis mängimiseks¹¹.

Lisaks on eelistatud vormingute hulka arvatud ka MOV. Tegemist on patentitud vorminguga, kuid on soovituslik erinevate institutsioonide poolt üle maailma, võimaldab faile säilitada tihendamata ning on laialdaselt kasutusel.

Aksepteeritavate vormingute hulka kuuluv MPEG-4 Part 14 on kadudega kokku pakkimist võimaldav vorming, mis hetkel laialdaselt kasutusel, aga pole patentitud. Vormingut kasutatakse näiteks teleasaadete edastamisel.

Eelistatud vormingud

Vorming ja laiend	Kooder	Standard
Materjal Exchange Format (.mxf) OP1a	JPEG 2000 kadudeta kompresseeritud	<i>SMPTE ST 377-1:2011, Material Exchange Format (MXF) File Format Specification:</i> http://standards.smpte.org/ <i>ISO/IEC 15444-1:2016, Information technology – JPEG 2000 image coding system: Core coding system:</i> http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=70018
QuickTime (.mov)	Kompresseerimata 4 : 2 : 2	<i>QuickTime File Format Specification:</i> https://developer.apple.com/library/content/documentation/QuickTime/QTFF/QTFFPreface/qtffPreface.html

Aksepteeritavad vormingud

Vorming ja laiend	Kooder	Standard
-------------------	--------	----------

¹¹ https://deepblue.lib.umich.edu/bitstream/handle/2027.42/83222/Video-Best_practices.pdf?sequence=1 (23.03.2017)

MPEG 4 Part 14 (.mp4)	AVC/H.264	ISO/IEC 14496-14:2003, Information technology -- Coding of audio-visual objects - Part 14: MP4 file format: http://www.iso.org/iso/catalogue_detail.htm?csnumber=38538 H.264: Advanced video coding for generic audiovisual services: https://www.itu.int/rec/T-REC-H.264-201610-I/en
Audio Video Interleaved Format (.avi)	JPEG 2000 DV-PAL AVC/H.264	AVI RIFF File Reference: https://msdn.microsoft.com/en-us/library/ms779636.aspx H.264: Advanced video coding for generic audiovisual services: https://www.itu.int/rec/T-REC-H.264-201610-I/en

Veebisaidid

Veebiarhiveerimiseks tuleb kasutada WARC vormingut. WARC vorming võimaldab koondada veebist korjatud informatsiooni (nt veebilehe) kokku ühte faili. WARC on edasiarendus Internet Archive ARC failivormingust, kuid võimaldab paremat tuge veebiarhiveerimisele ja juurdepääsule. WARC faili juures on ka metaandmed veebisaidi kohta.

Nimi	WARC (<i>Web ARChive file format</i>)
Failinime laiend	.warc
Standard	ISO 28500:2009 Information and documentation -- WARC file format: https://www.iso.org/standard/44717.html

Muud failid ja failivormingud

Väljaandja võib edastada ka muid trükieelseid algmaterjalide faile, mis on eelnenud lõpliku väljaande failile (nt Indesign'is, Photoshop'is jms loodud failid, heliteose lõplikule versioonile eelnenud versioonid jne). Failid tuleb Kirjastajaportaalis märkida tööfailideks ning need on kättesaadavad ainult kirjastajale, kuid mitte lugejale.

Juhul kui väljaandja edastatud fail ei ole pikaajaliseks säilitamiseks valiidne, võtame väljaandjaga ühendust ning palume saata sobiliku faili. Esmalt teavitame, et tuleks saata fail, mida me saame arhiveerida. Kui on eriline formaat, mida RR ei saa avada ega konverteerida säilitusvormingusse, palume saata pabereksemplar, et saaksime selle digiteerida¹².

¹² Otsus tehtud DIGARi koosolekul 11.04.2017.a

Faile, mis on hetkel digitaalarhiivis ja ei ole eelistatud failivormingus, hoiaime hinnanguliselt 2022. aastani migreerimata säilitusvormingusse. Kui toimub digitaalarhiivi sisu migreerimine uude arhiivi, konverteeritakse kõik failid soovituslikku ajakohasesse säilitusvormingusse.

Säilitusvorming on Rahvusraamatukogu loodud kriteeriumite põhjal valitud failivorming eesmärgiga säilitada väljaande sisu pikaajaliselt. Kui väljaandja poolt edastatud fail ei ole sellises vormingus, siis viiakse fail vastavasse vormingusse Rahvusraamatukogu poolt ehk toimub migreerimine. Kuna migreerimise käigus võivad mõned faili funktsionaalsused kaduma minna, siis on soovituslik faile edastada vastavalt Rahvusraamatukogu poolt esitatud nõuetele.