

Aineiston järjestäminen ja nimeäminen Musiikkiarkistossa

Documents related to the management of the Music Archive Finland's resources

1 "Aineiston järjestäminen ja nimeäminen Musiikkiarkistossa" (Organizing and naming of resources in the Music Archive Finland)

2 "Aineiston LTFS-nauhavarmennus Musiikkiarkistossa" (LTFS tape backup of resources in the Music Archive Finland)

3 "Aineiston hallinta Musiikkiarkistossa" (Management of resources in the Music Archive Finland).

This document introduces the naming practices of the Music Archive Finland's resources. All resources processed in the archive (file, object) must comply with the specifications of this document.

The resources of the Music Archive Finland are preserved in two entities: general collections and personal or community archives. The files are stored in a hierarchical directory structure according to the collection/archive, which corresponds to the structure of the original archive creation. Every file in the digital archive is given a unique filename that may contain only predefined printable US-ASCII characters.

In general collections, the file name starts with the collection ID, which is a string of letters. In personal and community archives, the file name starts with the archive ID, which is a running number.

The archive or collection ID is followed by either a media ID or a series ID. The previously used practice of creating file names was based on a media ID, where the material was organized by different media types. However, the new practice of creating file names is based on organizing material according to their content. It utilizes a unique series ID that is independent of the original media type.

After the media or serial ID, the file name includes a section that indicates the archival unit, followed by characters that distinguish the different archival unit groups within that archival unit.

If there are files that are related to the original file, or files that define the original file more precisely, additional sections of special identifiers are used in their file names. And if digital copies of different quality are to be created from the original file, these are also distinguished from the original copy by using additional identifiers.

Damaged digital objects can also be marked with specific identifiers. Furthermore, if the original material is removed from the archive or the life cycle of the file format ends, the special identifiers make it possible to keep records that guarantee the authenticity of the materials in the Skaala database managed by the Music Archive Finland.

For more details see the document "Aineiston hallinta Musiikkiarkistossa" (Management of resources in the Music Archive Finland).

Sisällys

| | |
|--|-----------|
| 1 Sisällys | 3 |
| 2 Johdanto | 5 |
| 3 Musiikkiarkiston kokoelmat ja arkistot | 5 |
| 3.1 Uusi ja vanha arkistonmuodostustapa | 6 |
| 3.2 Hakemisto- ja tiedostonimien tunnuksset ja nimiosiot | 6 |
| 4 Hakemisto- ja tiedostonimien rakenne: pää-, perus- ja tiedosto-osio | 6 |
| 4.1 Hakemistonimen rakenne | 6 |
| 4.1.1 Vanha arkistonmuodostustapa: yleiskokoelman mediakokoelmien hakemistonimien rakenne | 7 |
| 4.1.2 Vanha arkistonmuodostustapa: henkilö- ja yhteisöarkistojen hakemistonimien rakenne | 7 |
| 4.1.3 Uusi arkistonmuodostustapa: yleiskokoelman erikoiskokoelmien hakemistonimien rakenne | 7 |
| 4.1.4 Uusi arkistonmuodostustapa: henkilö- ja yhteisöarkistojen hakemistonimien rakenne | 8 |
| 4.2 Tiedostonimen rakenne | 9 |
| 5 Objektien järjestäminen Musiikkiarkiston järjestelmissä | 10 |
| 5.1 Kokoelma- ja arkistotunnukset | 10 |
| 5.1.1 Y: Yleiskokoelmat | 10 |
| 5.1.1.1 Yleiskokoelman mediakokoelmat | 10 |
| 5.1.1.2 Yleiskokoelman erikoiskokoelmat | 11 |
| 5.1.2 H: Henkilö- ja yhteisöarkistot | 12 |
| 5.2 Sarjatunnukset | 12 |
| 5.2.1 Sarjatunnukset (uusi arkistonmuodostustapa) | 12 |
| 5.2.2 Mediatunnukset sarjatunnuksina (vanha arkistonmuodostustapa) | 12 |
| 5.2.3 Mediatunnukset | 12 |
| 5.2.4 Arkistoyksiköt ja mediatunnukset | 13 |
| 5.2.4.1 Arkistoyksiköt ja mediatunnukset vanhassa arkistonmuodostustavassa | 13 |
| 5.2.4.2 Arkistoyksiköt ja mediatunnukset uudessa arkistonmuodostustavassa | 14 |
| 5.2.5 Osakokonaisuustunnus | 14 |
| 5.3 Kokoelma- tai arkistotunnuksen tiedosto-osio | 15 |
| 5.3.1 Numero-osio | 15 |
| 5.3.1.1 Osaobjekti- ja migraatio-osaobjektitunnus | 15 |
| 5.3.2 Vaihtoehtoinen digitointi -tunnus | 16 |
| 5.4 Kokoelma- tai arkistotunnuksen lisäosiot | 16 |
| 5.4.1 Lisätallennintunnus | 16 |
| 5.4.2 Lisätunnukset: liite-, litterointi ja indeksitiedostot | 18 |
| 5.4.2.1 Liitetunnus | 18 |
| 5.4.2.2 Litterointitunnus | 18 |
| 5.4.2.3 Indeksitunnus | 19 |
| 5.4.3 Käyttökopiotunnus (__kk + numero tai __kk0 + lisätunnus) | 19 |
| 5.4.3.1 __kk[1-9] | 19 |
| 5.4.3.2 __kk0 | 20 |
| 5.4.3.3 Sensuroitu käyttökopiotunnus (__skk[+ numero]) | 20 |
| 5.5 Syntysähköisen tai digitaalisessa muodossa vastaanotetun aineiston erikoistapaukset | 21 |
| 5.6 Erikoistunnukset | 22 |
| 5.6.1 Vauriotunnus: ---dmgd | 22 |
| 5.6.2 Migroitutunnus: ---migrated-<vvvkkpp>[-osaobjektitunnus] | 22 |
| 5.6.3 Poistetun objektin tunnus: ---deleted-<vvvkkpp> | 23 |
| 6 Liitteet | 24 |
| 6.1 Liite 1. Tiedostojen nimeämisesimerkkejä | 24 |
| 6.1.1 Vanha arkistonmuodostustapa | 24 |
| 6.1.1.1 Yleiskokoelmien mediakokoelmat | 24 |
| 6.1.1.2 Henkilö- ja yhteisöarkistot | 25 |

| | |
|---|----|
| 6.1.2 Uusi arkistonmuodostustapa..... | 25 |
| 6.1.2.1 Yleiskokoelmien erikoiskokoelmat | 26 |
| 6.1.2.2 Henkilö- ja yhteisöarkistot | 26 |
| 6.2 Liite 2. hakemisto- ja tiedostonimien osioiden säännölliset lausekkeet | 27 |
| 6.2.1 Henkilö- ja yhteisöarkistotunnus | 27 |
| 6.2.2 Yleiskokoelman erikoiskokoelmatunnus..... | 27 |
| 6.2.3 Sarjatunnus uudessa arkistonmuodostustavassa..... | 27 |
| 6.2.4 Sarjatunnus vanhassa arkistonmuodostustavassa (mediatunnus)..... | 27 |
| 6.2.5 Arkistoyksikkötunnus vanhassa arkistonmuodostustavassa (juokseva numero) | 27 |
| 6.2.6 Arkistoyksikkötunnus uudessa arkistonmuodostustavassa (juokseva numero + mediatunnus + juokseva numero) | 27 |
| 6.2.7 Osakokonaisuustunnus | 27 |
| 6.2.8 Numero-osio | 27 |
| 6.2.9 Osaobjektitunnus..... | 27 |
| 6.2.10 Migraatio-osaobjektitunnus | 27 |
| 6.2.11 Lisätallennintunnus | 27 |
| 6.2.12 Lisätunnukset | 27 |
| 6.2.12.1 Liitetunnus | 27 |
| 6.2.12.2 Litterointitunnus | 27 |
| 6.2.12.3 Indeksitunnus | 28 |
| 6.2.12.4 Liitenumero-osio | 28 |
| 6.2.13 käyttökopiotunnukset | 28 |
| 6.2.13.1 __kk0 | 28 |
| 6.2.13.2 __kk1–9 | 28 |
| 6.2.14 Erikoistunnukset | 28 |
| 6.2.14.1 Vauriotunnus: ---dmgd | 28 |
| 6.2.14.2 Migroititunnus: ---migrated-<vvvkkpp>[-osaobjektitunnus] | 28 |
| 6.2.14.3 Poistetun objektin tunnus: ---deleted-<vvvkkpp> | 28 |
| 6.3 Liite 3. Säännölliset lausekkeet eri kokoelmien ja arkistojen tiedostonimille..... | 29 |

Johdanto

Musiikkiarkiston aineiston hallintaan liittyvät dokumentit:

1 Aineiston järjestäminen ja nimeäminen Musiikkiarkistossa

2 Aineiston LTFS-nauhavarmennus Musiikkiarkistossa

3 Aineiston hallinta Musiikkiarkistossa

Tässä dokumentissa esitellään Musiikkiarkiston aineiston nimeämiskäytänteet. Aineistolla tarkoitetaan digitaalisessa muodossa olevia objekteja (tiedostoja), jollei toisin mainita.

Musiikkiarkistossa käsiteltävän aineiston tulee olla tämän dokumentin määritysten mukaista. Viitaukset pitkäaikaissäilytykseen (PAS), ks. dokumentti "Aineiston hallinta Musiikkiarkistossa". Viitaukset nauhavarmennukseen, ks. dokumentti "Aineiston LTFS-nauhavarmennus Musiikkiarkistossa".

Musiikkiarkiston kokoelmat ja arkistot

Musiikkiarkiston aineistot sijaitsevat kahdessa kokonaisuudessa: yleiskokoelmissa sekä henkilö- tai yhteisöarkistoissa. Yleiskokoelmat jakautuvat edelleen kahteen kokonaisuuteen: mediatunnuskokoelmiin ja erikoiskokoelmiin. Henkilö- ja yhteisöarkistot nimetään lähtökohtaisesti arkistonmuodostajan mukaan ja ne saavat tunnuksiksi juoksevan numeron.

Objektit (tiedostot) on tallennettu kokoelman tai arkiston mukaiseen hierarkkiseen hakemistorakenteeseen. Tiedostonimissä saa esiintyä vain tulostuvien US-ASCII-merkkien osajoukko. Sallittujen ASCII-merkkien desimaalikoodit

- 45: yhdysviiva "-"
- 46: piste "."
 - Sallittu vain yhden kerran ennen tiedostopäätettä
- 48–57: numerot 0–9
- 65–90: suuraakkoset A–Z
- 95: alaviiva "_"
- 97–122: pienaakkoset a–z

Ks. esimerkiksi <http://www.columbia.edu/kernit/ascii.html> ja <https://www.asciitable.com>.

Tiedostonimet eivät täten täysin vastaa kuvailevan metadatan järjestelmien tunnuksia (signum), esimerkiksi tiedoston "123__S-1-1__1Av1_1.mov" tunnus olisi "123 S.1.1:1Av1". Lisäksi kuvailevan metadatajärjestelmän tunnus liittyy lähes aina sarja- tai arkistoyksikkötasoon, jolloin tiedostonimestä pitää vielä poistaa numero-osio ja sitä mahdollisesti seuraavat osiot, jotta tunnukset täsmäisivät keskenään.

Jatkossa tiedostonimistä puhuttaessa sivuutetaan tiedostomuodon päätte muutoin, kuin käsitellessä useita eriformaattisia tiedostoja, joilla on – ja pitää olla – sama nimi. Tätä poikkeusta lukuun ottamatta tiedostonimet ovat Musiikkiarkistossa ainutlaatuisia objektien tunnuksia – joskin näissäkin tapauksissa tiedostonimi päätteineen muodostaa ainutlaatuisen merkkijonon.

Uusi ja vanha arkistonmuodostustapa

Digitaalisten objektien nimeämisessä käytetään sekä vanhaa että uutta arkistonmuodostustapaa. Uudessa arkistonmuodostustavassa aineisto ryhmitellään ensisijaisesti sisällön mukaan sarjoihin, joiden tunnukset muodostetaan kirjaimista tai numeroista. Sarjoihin sijoitettu aineisto tunnistetaan edelleen tarkemmalla tasolla (arkistoyksikkö ja siihen kuuluvat tiedostot) numeroita ja mediatunnuksia käyttäen. Uudessa arkistonmuodostustavassa yhden arkistoyksikön alle on mahdollista sijoittaa monipuolisesti sekä sama- että eriformaattisia aineistoja.

Vanha arkistonmuodostustapa rakentuu tallennusformaattien varaan, joille on annettu mediatunnukset, joita käytetään uuden arkistonmuodostustavan sarjatunnuksen tavoin. Tämän seurauksena vanhassa arkistonmuodostustavassa yksi arkistoyksikkö – kuten koko mediasarja/-kokoelma – koostuu pääasiallisesti vain yhdestä tallenneformaattista.

Hakemisto- ja tiedostonimien tunnukset ja nimiosiot

Kaikki nimeämissäännöt ovat yhtenäiset yleiskokoelmissa sekä henkilö- ja yhteisöarkistoissa, silloin kun kyseistä tunnusosiota käytetään. Esimerkiksi mediatunnukset ovat käytössä kaikissa kokoelmissa ja arkistoissa – joskin eri funktiossa – mutta henkilö- ja yhteisöarkistoissa ei missään tapauksessa ole käytössä yleiskokoelman erikoiskokoelmatunnuksia.

Kirjaimista koostuville tunnuksille on määritetty kirjainkoko suuri tai pieni. Kirjainkoot on valittu ainoastaan visuaalisiin perusteisiin eivätkä ne itsessään merkitse mitään. Tunnuksia ei myöskään voi sekoittaa toisiinsa kirjainkoon perusteella.

HUOM! Poikkeuksena edelliseen on osakokonaisuus- (suuraakkonen) ja osaobjektitunnukset (pienaakkonen), jotka molemmat ovat usein yhden kirjaimen mittaisia. Tunnukset ovat kuitenkin erotettavissa toisistaan niiden sijainnin perusteella hakemisto- tai tiedostonimessä.

Hakemisto- ja tiedostonimien rakenne: pää-, perus- ja tiedosto-osio

Tiedostonimi jakautuu kahteen osioon: hakemistonimi- ja tiedosto-osioon, jotka edelleen jakautuvat pienempiin osioihin. Hakemistonimi on lähtökohtaisesti arkistoyksikön tunnus, joka pitää sisällään yhden loogisen kokonaisuuden, esimerkiksi yhden CD-levyn, muusikkohaastattelun tai tapahtumatallenteen.

Hakemistonimen rakenne

Hakemistonimi koostuu pää- ja perusosiosta. Pääosio ilmaisee arkiston tai kokoelman, johon aineisto kuuluu. Perusosio ilmaisee sarja- ja arkistoyksikkötasot.

Pakolliset vähimmäisvaatimukset hakemistonimelle riippuvat arkistosta tai kokoelmasta, johon aineisto kuuluu. Alla esitetään hakemisto- ja tiedostonimien rakenne. Osiot esitellään tarkemmin myöhemmin.

HUOM! Kokoelmien ja arkistojen ylimmän tason hakemistonimi voi olla mikä tahansa, esimerkiksi "Y", "E", "H", "Y-arkisto" tai "H-arkisto". Ylimmän tason hakemistonimiä ei merkitä alihakemisto- eikä tiedostonimiin.

Vanha arkistonmuodostustapa: yleiskokoelman mediakokoelmien hakemistonimien rakenne

Hakemistonimi voi olla lyhimmillään kahden merkin mittainen: mediatunnus + numero.

Hakemistonimen rakenne:

- pääosio
 - **mediatunnus** (pakollinen)
- perusosio
 - **juokseva numero** (pakollinen)
 - *osakokonaisuustunnus* (valinnainen)

Esimerkiksi tunnuksen "Ck1" osiot ovat

- pääosio
 - sarja-/mediatunnus: Ck
 - C-kasettisarja
- perusosio
 - arkistoyksikkötunnus: 1
 - (ei osakokonaisuustunnusta)

Vanha arkistonmuodostustapa: henkilö- ja yhteisöarkistojen hakemistonimien rakenne

Hakemistonimi voi olla lyhimmillään kolmen merkin mittainen: arkistotunnus + mediatunnus + juokseva numero.

Hakemistonimen rakenne:

- pääosio
 - **arkistotunnus** (pakollinen)
 - **mediatunnus** (pakollinen)
- perusosio
 - **juokseva numero** (pakollinen)
 - *osakokonaisuustunnus* (valinnainen)

Esimerkiksi tunnuksen "21Md1A" osiot ovat

- pääosio
 - arkistotunnus: 21
 - mediatunnus: Md
 - MiniDisc-sarja
- perusosio
 - arkistoyksikkötunnus: 1
 - osakokonaisuustunnus: A
 - Osakokonaisuustunnuksia käytetään, jos yksittäinen media tai muu tallennekokonaisuus koostuu vähintään kahdesta eri kokonaisuudesta
 - Suuraakkosia käytetään juoksevan numeron tavoin

Uusi arkistonmuodostustapa: yleiskokoelman erikoiskokoelmien hakemistonimien rakenne

Hakemistonimen tunnuksen rakenne on lähes identtinen henkilö- ja yhteisöarkistojen rakenteen kanssa (ks. alla). Tunnukset erona on pääosion alku: yleiskokoelman erikoissarjoissa käytetään kokoelmatunnusta (kirjainmerkkijono) – henkilö- ja yhteisöarkistoissa käytetään arkistotunnusta (numero).

Hakemistonimen rakenne:

- pääosio
 - **kokoelmatunnus** (pakollinen)
 - **sarjatunnus erottimiseen** (pakollinen)
 - erotinmerkkijono: "___"

- perusosio
 - **juokseva numero** (pakollinen)
 - **mediatunnus** (pakollinen)
 - **juokseva numero** (pakollinen)
 - *osakokonaisuustunnus* (valinnainen)

Esimerkiksi tunnuksen ”Hst__1-2__1X1” osiot ovat

- pääosio
 - kokoelmatunnus: Hst
 - sarjatunnus erottimiseen: __1-2__
 - sarjaerotin: __
 - sarjatunnus: 1-2
 - ensimmäisen (pää)sarjan toinen (ala)sarja
 - sarjaerotin: __
- perusosio
 - arkistoyksikkötunnus: 1
 - mediatunnus: X
 - median numero: 1
 - (ei osakokonaisuustunnusta)

Erikoiskokoelmien sarjatunnukset koostuvat numeroista, esimerkiksi sarjatunnukset ”U-1-2” tai ”S-1-2” eivät ole mahdollisia.

Uusi arkistonmuodostustapa: henkilö- ja yhteisöarkistojen hakemistonimien rakenne

Hakemistonimen tunnuksen rakenne on lähes identtinen yleiskokoelmien erikoissarjojen rakenteen kanssa (ks. yllä). Tunnusten erona on pääosion alku: yleiskokoelman erikoissarjoissa käytetään kokoelmatunnusta (kirjainmerkkijono) – henkilö- ja yhteisöarkistoissa käytetään arkistotunnusta (numero).

Hakemistonimen rakenne:

- pääosio
 - **arkistotunnus** (pakollinen)
 - **sarjatunnus erottimiseen** (pakollinen)
 - erotinmerkkijono: ”__”
- perusosio
 - **juokseva numero** (pakollinen)
 - **mediatunnus** (pakollinen)
 - **juokseva numero** (pakollinen)
 - *osakokonaisuustunnus* (valinnainen)

Esimerkiksi tunnuksen ”1__S-1-2__1X1A” osiot ovat

- pääosio
 - arkistotunnus: 1
 - sarjatunnus erottimiseen: __S-1-2__
 - sarjaerotin: __
 - sarjatunnus: S-1-2
 - S-(pää)sarjan ensimmäisen (ala)sarjan toinen (ala)sarja
 - sarjaerotin: __
- perusosio
 - arkistoyksikkötunnus: 1
 - mediatunnus: X
 - median numero: 1
 - osakokonaisuustunnus: A
 - Osakokonaisuustunnuksia käytetään, jos yksittäinen media tai muu tallennekokonaisuus koostuu vähintään kahdesta eri kokonaisuudesta
 - Suuraakkosta käytetään juoksevan numeron tavoin

Henkilö- ja yhteisöarkistoissa sarjatunnukset voivat koostua sekä kirjaimista että numeroista tai vain numeroista, esimerkiksi sarjatunnukset "U-1-2" tai "1-1-2" ovat mahdollisia.

Tiedostonimen rakenne

Tiedosto-osion rakenne merkitsemissäonnöt ovat yhteiset kaikille arkistoille ja kokoelmille sekä uudessa että vanhassa arkistonmuodostustavassa. Tiedosto-osio lisätään hakemistonimen perään.

Tiedosto-osion rakenne:

- **numero-osio** (pakollinen)
 - *osaobjektitunnus* (valinnainen)
 - *migraatio-osaobjektitunnus* (valinnainen)
 - *lisätallennintunnus* (valinnainen)
 - *lisätunnus* (valinnainen)
 - *käyttökopiotunnus* (valinnainen)
 - *vaurio-, migrointi- tai poistetun objektin tunnus* (erikoistunnukset, valinnaisia)
 - Poistettujen objektien tunnukset ovat mahdollisesti tallessa vain Skaalan PAS-kirjanpidossa tai kuvailevan metatadan järjestelmissä

Esimerkiksi hakemistoon "1__1-2__1X1" voi olla talletettu tiedostoja, joiden tiedosto-osiot ovat

- _1
- _1_liite1
 - Liitetiedosto, joka liittyy samassa hakemistossa sijaitsevaan objektiin, jonka tiedosto-osio on "_1"
- _2a, _2b
 - Osaobjektitunnuksia käytetään, jos yksi objekti koostuu vähintään kahdesta eri tiedostosta
 - Pienaakkosta käytetään juoksevan numeron tavoin
- _2_liite1
 - Liitetiedosto, joka liittyy samassa hakemistossa sijaitsevaan objektiin, jonka tiedosto-osio on "_2"
 - Liitetiedostojen numerointi alkaa aina yhdestä, jos liitetosiota edeltävä osio on muuttunut
- _3

Esimerkit täysistä tiedostonimistä alempana.

Uuden arkistonmuodostusjärjestelmän mukaisen tiedostonimen rakenne kokonaisuudessaan:

- pääosio
 - **arkistotunnus** (pakollinen)
 - **sarjatunnus erottimineen** (pakollinen)
- perusosio
 - **juokseva numero** (pakollinen)
 - **mediatunnus** (pakollinen)
 - **juokseva numero** (pakollinen)
 - *osakokonaisuustunnus* (valinnainen)
- tiedosto-osio
 - **numero-osio** (pakollinen)
 - *osaobjektitunnus* (valinnainen)
 - *migraatio-osaobjektitunnus* (valinnainen)
 - *lisätallennintunnus* (valinnainen)
 - *lisätunnus* (valinnainen)
 - *käyttökopiotunnus* (valinnainen)
 - *vaurio-, migrointi- tai poistetun objektin tunnus* (erikoistunnukset, valinnaisia)
 - Poistettujen objektien tunnukset ovat mahdollisesti tallessa vain Skaalan PAS-kirjanpidossa tai kuvailevan metatadan järjestelmissä

Esimerkiksi hakemisto "1__1-2__1X1" voi sisältää tiedostot

- 1__1-2__1X1_1
- 1__1-2__1X1_1_liite1

- 1__1-2__1X1_2a ja 1__1-2__1X1_2b
- 1__1-2__1X1_2_liite1
- 1__1-2__1X1_3

Objektien järjestäminen Musiikkiarkiston järjestelmissä

Objektit on järjestetty hakemistoittain arkistonmuodostusta mukailevaan hierarkkiseen rakenteeseen. Hakemistoissa saa olla vain tiedostoja, joiden nimien alku täsmää hakemistonimen kanssa. Esimerkiksi hakemistossa "1Ck1" ei saa olla tiedostoja, joiden nimi alkaa merkkijonolla

- 1Ck1A
- 1Ck2
- 1Md1
- 2Ck1
- Hst__1 jne.

On kuitenkin mahdollista, että hakemistossa 1Ck1 on esimerkiksi alihakemistot "1Ck1A", "1Ck1B", "1Ck1C" ja "1Ck1D", joihin on sääntöjen mukaisesti sijoitettu tiedostot nimeltään "1Ck1A_1", "1Ck1B_1" ja niin edelleen. Alihakemistojen nimien alkujen on täsmättävä ylahakemistojen nimien alkujen kanssa. (Poikkeuksena ylimmän tason hakemistot, kuten mainittu ja myöhemmin esitetään.)

HUOM! Sääntöjä sovelletaan ensisijaisesti arkistoaineiston tunnuksiin, jotka annetaan arkistonmuodostuksen tai luetteloinnin yhteydessä. Digitoitujen aineistojen tiedostonimien tulee vastata arkistoaineistojen tunnuksia. On hyvin mahdollista, että sähköisessä arkistossa esimerkiksi hakemistossa "1Av" on vain hakemistot "1Av3" ja "1Av5". Tämä tarkoittaa, että vain arkistoyksiköt "1Av3" ja "1Av5" on digitoitu, ja että arkistoon on talletettu digitoitujen arkistoyksiköiden lisäksi vähintään arkistoyksiköt "1Av1", "1Av2" ja "1Av4". Tunnusten juoksevien numeroiden on oltava katkeamattomia ja ne alkavat aina numerosta yksi kussakin osiossa (sarjassa, arkistoyksikössä, arkistoyksikön numero-osiossa, lisätunnusosiossa jne.). Lisäksi on kuitenkin huomattava, että arkistojen juoksevat numeroinnit eivät kaikissa tapauksissa ole täydellisiä. Tällöin aineistokokonaisuuden metatiedoista on käytävä ilmi, onko arkistoyksiköitä esimerkiksi poistettu arkistosta vai onko kyseessä ollut luetteloinnin aikainen numerointivirhe.

Kokoelma- ja arkistotunnukset

Tästä alkaa hakemistonimen osuus. Yleiskokoelmissa hakemistonimi alkaa aina isolla kirjaimella, henkilö- ja yhteisöarkistoissa numerolla 1–9. Tiedostonimeä koskee samat säännöt, sillä tiedostonimeen lisätään aina hakemistonimi. Etunollia, joilla saataisiin jokin tunnusosio aina viiden numeron mittaiseksi, ei käytetä.

Y: Yleiskokoelmat

Yleiskokoelman tunnus on "Y". Yleiskokoelma jakautuu kahteen kokoelmaan: mediatunnuskoelmiin ja erikoiskokoelmiin "Y-E". Digitaalinen objekti tai hakemisto liittyy aina johonkin yleiskokoelmaan, jos sen nimi alkaa kirjaimella.

HUOM! Yleiskokoelmatunnus "Y", "Y-E" esiintyy vain ylätason hakemistonimessä.

Yleiskokoelman mediakokoelmat

Mediakokoelmatunnus (mediatunnus) on yhdestä viiteen merkkiä pitkä kirjainmerkkijono, joka alkaa isolla kirjaimella ja voi päättyä isoon tai pieneen kirjaimeseen. Mediatunnukset esitetään tarkemmin alempana.

Yleiskokoelman mediakokoelmat (mediasarjat) ovat ryhmitelty formaateittain: samankaltaiset mediat ovat yleensä samankokoisia ja -muotoisia, joten niitä on luontevaa käsitellä yhtenä kokonaisuutena ja sijoittaa arkistotiloissa lähekkäin.

Hakemisto- ja tiedostonimet, ylätasoinen hakemistonimeä lukuun ottamatta, alkavat mediatunnuksella (ks. alla) ja ovat täten muotoa {mediatunnus}{juokseva numero}{...}.

Hakemistonimet esimerkiksi

- Av1
- Av2
 - Av2 ei ole sallittu ilman Av1:tä
- Ck21A, Ck21B, Ck21C
 - Ck21A ei ole sallittu ilman Ck21B:tä
 - Ck21B ei ole sallittu ilman Ck21A:ta
 - Ck21C ei ole sallittu ilman Ck21A:ta ja Ck21B:tä
- Md321

Tiedostonimet esimerkiksi

- Av1_1
- Av2_1a, Av2_1b
 - Av2_1a ei ole sallittu ilman Av2_1b:tä
 - Av2_1b ei ole sallittu ilman Av2_1a:ta
- Ck21A_1, Ck21B_1-1, Ck21B_1-2
 - Ck21B_1-1 ei ole sallittu ilman Ck21B_1-2:ta
 - Ck21B_1-2 ei ole sallittu ilman Ck21B_1-1:tä
 - Hakemisto Ck21B ei ole sallittu ilman Ck21A:ta eikä hakemisto Ck21A ei ole sallittu ilman Ck21B:tä
- Md321_1-1, Md321_1-2, Md321_1-3
 - Md321_1-1 ei ole sallittu ilman Md321_1-2:ta
 - Md321_1-2 ei ole sallittu ilman Md321_1-1:tä
 - Md321_1-3 ei ole sallittu ilman Md321_1-1:tä ja Md321_1-2:ta

Sallittuihin hakemisto- ja tiedostonimiin liittyen ks. huomio yllä arkistoaineiston tunnuksista ja aineiston tapauskohtaisen digitoimisen vaikutus digitaalisesti objektien sijoittamisesta sähköiseen arkistoon.

Yleiskokoelman erikoiskokoelmat

Erikoiskokoelmatunnus on yhdestä viiteen merkkiä pitkä kirjainmerkkijono, joka alkaa isolla kirjaimella ja voi päättyä isoon tai pieneen kirjaimeseen. Kirjaintunnusta seuraa merkkijono "___". Tällä hetkellä (20221220) on neljä erikoiskokoelmaa

- Hst (Y:E:Hst): Haastattelut
- Ker (Y:E:Ker): Keruut
- Smj (Y:E:Smj): Sisällön mukaan järjestetyt aineistot
 - "Sisältö" tarkoittaa tässä yhteydessä hyvin tarkkaan rajattua kokonaisuutta, joka on mielekästä pitää yhdessä. Sarjassa on esimerkiksi radio-ohjelmasarja Jee Jee Jee – Suomalaisen rockin neljä vuosikymmentä ja musiikkilehtien skannauksia
- Tdt (Y:E:Tdt): Levy-yhtiöiden tiedotteet

Ylätasoinen hakemistonimeä lukuun ottamatta hakemisto- ja tiedostonimet alkavat erikoiskokoelmatunnuksella, jonka jälkeen tulee sarjaerotin "___", ja ovat täten muotoa {erikoiskokoelmatunnus}___{sarjatunnus}___, esimerkiksi

- Hst___1-1-1___
- Smj___2-2___

HUOM! Käsikirjaston kirjat ("Y-K") on luetteloidessa nimetty tekijöiden ja vuosilukujen mukaan, vaikka ovat mediatunnuskokoelmaa. Käsikirjaston sijaintitunnukset "K1" (yleisteokset), "K2" (biografiat), "K3" (diskografiat) ja "K4" (tekniikka).

HUOM! On olemassa myös toinen tapa, esimerkiksi

- hakemisto MusA__Y-E-Ker__5__1Smj1 sisältää tiedostot
 - Ker__5__1Smj1_IMG_8835.JPG
 - Ker__5__1Smj1_IMG_8836.JPG
- hakemisto __Y-E-Ker__4-1-2__ sisältää tiedostot
 - __Y-E-Ker__4-1-2__1Smj1_Juhon-paikat.docx
 - __Y-E-Ker__4-1-2__1Smj1_Tuulin-ajatuksia.docx
- hakemisto MusA__Y-E-Hst__3-1__ sisältää tiedostot
 - MusA__Y-E-Hst__3-1__17Y1_1.doc
 - MusA__Y-E-Hst__3-1__17Y1_1.pdf
 - MusA__Y-E-Hst__3-1__17Y1_1.txt

Objektien nimissä sallitaan myös osiot "MusA__Y-E-" ja "__Y-E-", tai korjataan tilanne jossain sopivassa varmuuskopiointivaiheessa. TODO! Korjataan siten, että objektien nimet alkavat erikoiskoelman kirjainmerkkitunnuksella: Hst, Ker, Smj tai Tdt.

H: Henkilö- ja yhteisöarkistot

Henkilö- ja yhteisöarkistojen tunnus on numero. Objekti – tai hakemisto – liittyy aina johonkin henkilö- tai yhteisöarkistoon, jos sen nimi alkaa numerolla.

Vanhassa arkistonmuodostustavassa pääosioon kuuluu myös mediatunnus, joka seuraa välittömästi arkistotunnusta. Uudessa arkistonmuodostustavassa pääosioon kuuluu myös sarjatunnus erottimiseen.

Sarjatunnukset

Päätason kokoelma- tai arkistotunnuksen jälkeen merkitään sarjatunnus.

Sarjatunnukset (uusi arkistonmuodostustapa)

Uudessa arkistonmuodostustavassa sarjatunnus on kirjain- tai numeromerkintä, joka erotetaan sarjatunnuserottimella ”__” sitä ympäröivistä osioista. Sarjatunnuksen pituutta ei ole rajoitettu: sarjatunnus vastaa arkistonmuodostuksen yhteydessä luotuja sarjatunnuksia.

Esimerkiksi hakemisto

- 1__S-1__ sisältää arkiston numero 1 sarjan S.1
- 21__3-2-1__ sisältää arkiston numero 21 sarjan 3.2.1

Mediatunnukset sarjatunnuksina (vanha arkistonmuodostustapa)

Vanhassa arkistonmuodostustavassa mediatunnuksia käytetään sarjatunnuksina, ja ne liitetään suoraan kiinni niitä ympäröiviin osioihin käyttämättä erotinmerkkiä.

HUOM! Mediatunnus kuuluu hakemistonimen pääosioon vanhassa arkistonmuodostustavassa. Uudessa arkistonmuodostustavassa mediatunnus kuuluu hakemistonimen perusosioon: tunnukset ovat identtiset molemmissa yhteyksissä. Mediatunnusten käyttö uudessa ja vanhassa arkistonmuodostustavassa käsitellään tarkemmin alempana.

Mediatunnukset

Mediatunnus on yhdestä viiteen merkkiä pitkä kirjainmerkkijono, joka alkaa isolla kirjaimella ja voi päättyä isoon tai pieneen kirjaimeseen.

- Ak: asiakirja
- Av: audiovisuaalinen aineisto
- Cd: CD-levy
 - Käytetään, jos CD-levy on musiikkijulkaisu. Itse poltettujen datalevyjen sisällölle annetaan aineiston mukaiset tunnukset
- Ck: C-kasetti
- Dat: Dat-nauha

- Dvd: DVD-levy
 - Käytetään, jos DVD-levy on audiovisuaalinen julkaisu. Itse poltettujen datalevyjen sisällölle annetaan aineiston mukaiset tunnukset
- Ep: EP-levy
 - Selkeästi LP-levyä pienempi kotimainen levy
- EpU: EP-levy, ulkomainen
 - Selkeästi LP-levyä pienempi ulkomainen levy
- Ju: juliste
- K: kirja
- Kn: avokelanauha
- L: lehti
- Le: LP-levy
 - EP:tä tai singleä isompi kotimainen levy, myös savikiekot tms.
 - Pikalevyt on merkitty tunnukseksi X
- LeU: LP-levy, ulkomainen
 - EP:tä tai singleä isompi ulkomainen levy, myös savikiekot tms.
 - Pikalevyt on merkitty tunnukseksi X
- Ll: lehtileike
- Md: MiniDisc-levy
- Mk: muu kasetti
 - Mini-, mikro, Elcaset, 8-track jne.
- Mm: multimedia
 - Käytetään, jos kyseessä on esimerkiksi multimedia-aineisto tai ohjelmisto, joka koostuu useista tiedostoista
- Smj: sisällön mukaan järjestetty tai nimetty
 - *HUOM! Käytössä myös yleiskokoelman erikoissarjatunnukseksi*
- Vk: valokuva
- Www: internetsivusto/-sivu
- X: äänitallenne
 - Geneerinen tunnus
 - Äänitallenne, jolle ei ole annettu median mukaista tunnusta
- Y: muu kuva tai teksti
 - Geneerinen tunnus
 - Kuva, kartta, kirje tms. dokumentti, jolle ei ole annettu median tai tarkoituksen mukaista tunnusta
- Z: määrittelemätön
 - Käytetään, jos alkuperäismedia tai sen tarkoitus on epäselvä. Kuitenkin esimerkiksi haastattelukokoelmien haastattelut merkitään X:ksi, jos ne ovat äänitallenteita ja Av:ksi, jos ne ovat videotallenteita

Arkistoyksiköt ja mediatunnukset

Arkistoyksikkö on tapauskohtaisesti määritetty yhtenäinen looginen kokonaisuus, esimerkiksi

- yksi C-kasetti, jolla on yksi haastattelu
- 5 C-kasettia, joissa on yhtenäinen sisältö
- yksi tapahtumatallenne, joka voi sisältää useita tallennusformaatteja, tai
- hyvinkin laaja kokonaisuus, jota ei ole mielekästä luetteloida tarkemmalla tasolla.

Arkistonmuodostusta tai digitointia suoritettaessa on aina tapauskohtaisesti ratkaistava, kuinka suuri tai pieni kokonaisuus ryhmitellään sarjaksi, ja missä vaiheessa aineisto on luontevinta määritellä arkistoyksiköiksi.

Arkistoyksiköt ja mediatunnukset vanhassa arkistonmuodostustavassa

Vanhassa arkistonmuodostustavassa aineisto on ryhmitelty (sarjoitettu) tallennusformaattien mukaan. Mediasarjoissa ja -kokoelmissa mediatunnuksen jälkeen tulee juokseva numero, joka ilmaisee kuinka mones arkistoyksikkö/media on kyseessä, esimerkiksi hakemisto

- 1Ck sisältää arkiston numero 1 C-kasetit, ja
 - 1Ck1 sisältää C-kasettisarjan ensimmäisen C-kasetin/arkistoyksikön

- 1Ck2 sisältää C-kasetisarjan toisen C-kasetin/arkistoyksikön jne.
- 21Md sisältää arkiston numero 21 MiniDisc-levyt, ja
 - 21Md1 sisältää MiniDisc-sarjan ensimmäisen MiniDisc-levyn/arkistoyksikön
 - 21Md2 sisältää MiniDisc-sarjan toisen MiniDisc-levyn/arkistoyksikön jne.

Vanha arkistonmuodostustapa on aina käytössä yleiskokoelman mediakokoelmissa, ja voi olla käytössä henkilö- ja yhteisöarkistoissa.

Arkistoyksiköt ja mediatunnukset uudessa arkistonmuodostustavassa

Uudessa arkistonmuodostustavassa sarjatunnus korvaa mediatunnuksen sarjan ilmaisijana. Sarjatunnuksen erottimen jälkeen merkitään arkistoyksikön ilmaiseva juokseva numero, joka erottaa sarjan arkistoyksiköt toisistaan. Arkistoyksikköä ilmaisevan numeron jälkeisillä mediatunnuksilla ja niitä seuraavilla numeromerkinnöillä ryhmitellään arkistoyksikön sisältämät objektit tapauskohtaisesti, esimerkiksi arkistoyksikössä "1__S-1__1"

- 1__S-1__1Vk1 sisältää arkiston numero 1 sarjan S.1 ensimmäisen arkistoyksikön ensimmäisen valokuvan/valokuvaryhmän
- 1__S-1__1Vk2 sisältää arkiston numero 1 sarjan S.1 ensimmäisen arkistoyksikön toisen valokuvan/valokuvaryhmän
- 1__S-1__1Av1 sisältää arkiston numero 1 sarjan S.1 ensimmäisen arkistoyksikön ensimmäisen videotallenteen/videotallenneryhmän
 - Jokaisen uuden mediatunnuksen juokseva numerointi alkaa yhdestä

Edellisessä esimerkissä "1__S-1__1Vk1", "1__S-1__1Vk2" ja "1__S-1__1Av1" ovat hakemistoja, mutta sähköisen aineiston järjestäminen on aina tapauskohtaista. Jos esimerkiksi kokonaisuus "1__S-1__1Av1" sisältää vain yhden videotiedoston "1__S-1__1Av1_1", on järkevää sijoittaa se suoraan ylemmän tason hakemistoon "1__S-1__1" valokuvahakemistojen "1__S-1__1Vk1" ja "1__S-1__1Vk2" rinnalle.

Uusi arkistonmuodostustapa on aina käytössä yleiskokoelman erikoiskokoelmissa, ja voi olla käytössä henkilö- ja yhteisöarkistoissa.

HUOM! 437__U-3__Vk: Mediatunnus ilman edeltävää arkistoyksikkönumeroa. Voidaan siirtää mediatunnuksen jälkeinen numero arkistoyksikkötunnukseksi tai sallitaan vanhakantainen tapa, jossa mediatunnusta seuraa arkistoyksikön numerotunnus, vaikka käytössä olisi uusi arkistonmuodostutapa.

Osakokonaisuustunnus

Osakokonaisuustunnus on suuri kirjain.

HUOM! Ainakin arkistossa 34 on kaksi isoa kirjainta merkitsemässä kelanauhan alku- ja loppunauhoja väritietoineen: GL: Green/Gold Lead(?); GR: Green/Gold Rear(?); RL: Red Lead(?); RR: Red Rear(?).

Jos yksi media on jaettavissa vähintään kahteen selkeään arkistoyksikköön – kokonaisuuteen –, lisätään juoksevan medianumeron perään osakokonaisuustunnus. Osakokonaisuustunnusta ei käytetä, jos osakokonaisuuksia ei ole vähintään kahta, esimerkiksi tunnuksen "A" käyttäminen ilman tunnusta "B" ei sallita, eikä tunnusta "B" sallita ilman tunnusta "A".

Vanhassa arkistonmuodostustavassa esimerkiksi hakemistot

- 1Ck1A ja 1Ck1B
 - Yhdessä C-kasetissa on kaksi arkistoyksikköä, esimerkiksi kahden eri henkilön haastattelut
- 21Av1A, 21Av1B ja 21Av1C
 - Yhdessä videokasetissa on esimerkiksi kolme eri arkistoyksikköä, esimerkiksi kolmen eri henkilön haastattelut

Uudessa arkistonmuodostustavassa vastaavasti esimerkiksi hakemistot

- 1__1-1__1Ck1A ja 1__1-1__1Ck1B
- 21__1-2__1Av1A, 21__1-2__1Av1B ja 21__1-2__1Av1C

Juokseva medianumero on siirretty uudessa arkistonmuodostustavassa mediatunnuksen eteen. Tämän lisäksi myös mediatunnuksen jälkeen merkitään juokseva numero, mikä mahdollistaa useiden saman tai eri formaattisten aineistojen liittämisen ja ryhmittelyn arkistoyksikön sisällä. Uusi ar-

kistonmuodostustapa mahdollistaa täten monipuolisesti hyvinkin laajojen aineistojen ryhmittelyn arkistoyksikön sisälle.

Tähän päättyy hakemistonimiosio (arkistoyksikön tunnus/signum), joka päättyy aina numeroon tai isoon kirjaimeseen.

Kokoelma- tai arkistotunnuksen tiedosto-osio

Tästä alkaa tiedostonimien osuus. Lähtökohtaisesti alla olevia osioita ei ole hakemistojen nimissä. On kuitenkin mahdollista, että esimerkiksi suuri määrä liitetiedostoja tietyn arkistoyksikön hakemistossa on sijoitettu yhteen alihakemistoon, jolloin sähköisen arkistoaineiston rakenteen hahmottaminen helpottuu.

Numero-osio

Hakemistonimiosion lisäksi objekteille ainoa pakollinen osio on numero-osio.

Numero-osio alkaa alaviivalla ”_”, jonka lisäksi sitä seuraava numero on pakollinen. Numeron jälkeen voi lisäksi esiintyä yhdysviiva ”-”, jonka jälkeen tulee esiintyä numero. Tiedostonimet esimerkiksi

- 1Ck1_1
- 1Ck2_1-1, 1Ck2_1-2, 1Ck2_1-3
 - 1Ck2_1-1 ei ole sallittu ilman 1Ck2_1-2:ta
 - 1Ck2_1-2 ei ole sallittu ilman 1Ck2_1-1:tä
 - 1Ck2_1-3 ei ole sallittu ilman 1Ck2_1-1:tä ja 1Ck2_1-2:ta

Osaobjekti- ja migraatio-osaobjektitunnus

Osaobjektitunnus on pieni kirjain, jota käytetään juoksevan numeron tavoin: a vastaa numeroa 1, b numeroa 2, c numeroa 3 ja niin edelleen.

Migraatio-osaobjektitunnus on pieni kirjain, jota edeltää yhdysviiva ”-”, ja jota käytetään juoksevan numeron tavoin: a vastaa numeroa 1, b numeroa 2, c numeroa 3 ja niin edelleen.

HUOM! Migraatio-osaobjektitunnusta edeltää yhdysviiva, jotta sitä ei ole mahdollista sekoittaa osaobjektitunnukseen, ks. "Migraatio-osaobjektitunnus: ---migrated-<vvvkkpp>[-osaobjektitunnus]". Täten on myös mahdollista migroida alempana olevan esimerkin objekti 445__H-3__1Y1_11 ilman, että korvaava objekti sekoittuu objektin 445__H-3__1Y1_11a kanssa.

HUOM! Migraatio-osaobjektitunnus on erikoistunnus, joka liitetään korvatun tiedostonimen perään, ja sitä voi seurata vain tiedostopääte edeltävine pisteineen. Migraatio-osaobjektitunnus on korvaavan objektin tunnus osa, joka liitetään numero-osion ja vaihtoehtoisen osaobjektitunnuksen perään.

Osaobjektitunnusta käytetään tilanteissa, joissa yhtä loogista kokonaisuutta tai objektia vastaa vähintään kaksi tiedostoa. Tässä yhteydessä loogisella kokonaisuudella tarkoitetaan arkistoyksikköä tarkempaa kokonaisuutta – yhtä objektia – joskin vain yksi objekti voi muodostaa arkistoyksikön, esimerkiksi

- juliste on skannattu useammassa osassa
 - Yhdistämällä kuvat saadaan kuva koko julisteesta
- pitkä tapahtumatallenne jakautuu kahdelle tai useammalle muistikortille tai tiedostoon
 - Tiedostoista on mahdollista luoda yhtenäinen käyttökopio yhdistämällä tiedostot
- ääni- tai videotallenne on digitoitu useaan kertaan
 - Pikalevyjä on saatettu digitoida käyttäen useita eri neuloja
 - Kelanauhoja on saatettu digitoida eri nauhureilla
 - Videokasetteja on saatettu digitoida eri nauhureilla
- objekti on migroitu kahteen tai useampaan kertaan, kahteen tai useampaan tiedostomuotoon
- leikekirjojen sivuja on digitoitu useaan kertaan, jos sivun kuvat tai leikkeet peittävät toisiaan
 - Jokainen objekti pyritään digitoimaan taittamalla sitä peittävät objektit pois edestä
 - Tiedostoista ei ole mahdollista koostaa yhtenäistä leikekirjan sivua, jossa kaikki sisältö on samanaikaisesti näkyvissä

Esimerkiksi tiedostot

- 1Ju1_1a, 1Ju1_1b ja 1Ju1_1c ovat yhden julisteen (osa)skannauksia
- 2X1_1a ja 2X1_1b voivat olla esimerkiksi
 - saman äänitallenteen eri tavalla tehtyjä digitointia
 - pitkän (syntysähköisen) äänitallenteen kaksi tiedostoa, jotka tallennin on jakanut useaan wav-tiedostoon tiedostokokorajoitteen (2 Gt) vuoksi

HUOM! On mahdollista, että leikekirjan kokonainen sivu on digitoitu ja numeroitu ilman osaobjektitunnusta, ja joku sivun yksittäinen objekti on digitoitu käyttäen osaobjektitunnusta. Tai että osaobjektitunnusta käyttäen on digitoitu eri objekteja eri laatuvaatimuksin.

- Esimerkiksi arkistossa 445
 - 445__H-3__1Y1_11: asiakirjalaatuinen digitointi, 600 dpi
 - 445__H-3__1Y1_11a: valokuvaalaatuinen digitointi, 900 dpi
 - *HUOM! Vain yksi osaobjektitunnus (a) on käytössä valokuvaalaatuksessa digitoinnissa. Asiakirjalaatuinen digitointi on jätetty merkitsemättä osaobjektitunnuksella*
 - 445 __H-3__1Y1_112a – 445 __H-3__1Y1_112e: asiakirjalaatuiset digitoinnit, 600 dpi
 - 445 __H-3__1Y1_112f: valokuvaalaatuinen digitointi, 900 dpi
 - 445__H-3__1Y2_20a ja 445__H-3__1Y2_20b: asiakirjalaatuiset digitoinnit, 600 dpi
 - 445__H-3__1Y2_20c: valokuvaalaatuinen digitointi, 900 dpi
 - 445__H-3__1Y2_116: asiakirjalaatuinen digitointi, 600 dpi
 - 445__H-3__1Y2_116a: valokuvaalaatuinen digitointi, 900 dpi
 - *HUOM! Vain yksi osaobjektitunnus (a) on käytössä valokuvaalaatuksessa digitoinnissa. Asiakirjalaatuinen digitointi on jätetty merkitsemättä osaobjektitunnuksella*

Vaihtoehtoinen digitointi -tunnus

Osaobjektitunnus on ensisijainen tapa ilmaista vaihtoehtoiset digitoinnit samasta aineistoista, mutta se soveltuu parhaiten tilanteisiin, joissa digitoinnit tehdään suunnitellusti peräkkäin, jolloin myös ensimmäiselle digitoinnille voidaan antaa osaobjektitunnus "a".

Mikäli aikaisemmin digitoidusta aineistosta tehdään uusi digitointi – eikä osaobjektitunnusta ole mahdollista tai mielekäästä käyttää – annetaan uusille digitoinneille vaihtoehtoisia digitointia merkitsevä tunnus: "-dig{nro}".

Aikaisempaa (ensimmäistä) digitointia ei merkitä: pienin käytettävä numero tunnuksessa on 2. Mikäli vaihtoehtoisia digitointeja tehdään useita, kasvatetaan tunnuksen numero-osiota vastaavasti.

Tapauskohtaisesti on päätettävä ja merkittävä, miten uusi digitointi suhtautuu vanhaan; onko mahdollista liittää vaihtoehtoisen digitoinnin tunnus vastaamaan olemassa olevia digitaalisia objekteja vai menevätkö tiedostojen sisällöt limittäin toistensa kanssa, esimerkiksi

- objekti 1Ep1_1-dig2
 - korvaa tai vastaa objektia 1Ep1_1
- objekti 1Ep2_1a-dig2
 - korvaa tai vastaa objektia 1Ep2_1a
- objekti 2Le1-dig2 voi korvata tai vastata olemassa olevia objekteja
 - 2Le1_1, 2Le1_2, 2Le1_3, jne.

"Korvaaminen" yllä tarkoittaa ensisijaisesti alkuperäisen aineiston täydentämistä erilaatuksilla digitoinneilla: aineistosta luodaan useita erilaisia kopioita, joita käyttäen on mahdollista luoda mahdollisimman laadukas rekonstruktio alkuperäisestä sisällöstä. Vaihtoehtoisesti olemassa olevat digitoinnit voidaan täysin korvata uusilla objekteilla, ks. "Migrointitunnus: ---migrated-<vvvkkpp>[-osaobjektitunnus]" ja "Poistetun objektin tunnus: ---deleted-<vvvkkpp>".

Kokoelma- tai arkistotunnuksen lisäosiot

Lisätunnukselliset tiedostot voidaan liittää arkistoitaviin objekteihin tai abstrakteihin kokonaisuuksiin kuten sarja- tai arkistoyksikkötasoihin.

Lisätallennintunnus

Lisätallennintunnus koostuu merkkijonosta "-cam" tai "-audio", jonka perään liitetään juokseva numero, jonka perään liitetään numero-osio.

*HUOM! Tilanteesta riippuen juokseva numero alkaa numerosta 1 tai 2. Numerointikäytäntö esitel-
lään alempana esimerkkien yhteydessä.*

Lisätallennintunnuksia käytetään tilanteissa, joissa esimerkiksi tapahtuma on tallennettu kahdella tai useammalla videokameralla tai äänitallentimella. Täten samanaikaisen tallentamisen seurauksena syntyneet tiedostot on luontevaa merkitä sisällöiltään toisiaan vastaaviksi. Tilanteesta riip-
puen on mahdollista käyttää joko lisätallennintunnusta tai uutta medianumeroa; lisätallennintunnus ilmaisee heti, että tapahtumatallenteesta on toinenkin kamerakulma tai äänitiedosto saatavissa, mutta tunnuksen puuttuminen ei takaa, ettei muita sisällöllisesti vastaavia tallenteita olisi saatavilla esimerkiksi muissa arkistoyksiköissä.

Lisätallennintunnuksia voidaan käyttää sekä vanhassa että uudessa arkistonmuodostustavassa. Tällä hetkellä (20200228) ne ovat käytössä vain uudessa arkistonmuodostustavassa. Vanhassa ar-
kistonmuodostustavassa ne voivat olla joissain tapauksissa ainoa vaihtoehto aineiston ryhmittele-
miseksi, sillä mediatunnusten mukainen ryhmittely arkistoyksikön sisällä on mahdollista vain uu-
dessa arkistonmuodostustavassa.

Poistamalla tiedostonimestä lisätallennintunnus saadaan kokonaisuus, johon objekti liittyy. Esi-
merkiksi tiedosto "1Av1_1-cam2_1" vastaa sisällöltään tiedostoa tai kokonaisuutta "1Av1_1". Ko-
konaisuus "1Av1_1" voi olla yksittäinen tiedosto tai esimerkiksi kahden tai useamman tiedoston
muodostama kokonaisuus

- 1Av1_1a ja 1Av1_1b tai
- 1Av1_1-1, 1Av1_1-2, 1Av1_1-3

HUOM! Ensisijaiseksi kameraksi valittua kameraa ei merkitä tunnuksella "-cam1".

Jos tapahtumasta on videotallenteiden lisäksi erikseen äänitettyä materiaalia, käytetään äänitie-
dostojen nimissä lisätallennintunnusta "-audio". Äänitallenteet merkitään lisätallennintunnuksella,
jos tapahtumasta on vähintään yksi videotallenne ja lisäksi vähintään yksi äänitallenne, esimerkiksi
videotallenteen "21__S-1__1Av1_1" sisältöä voi vastata äänitiedosto

- 21__S-1__1X1_1 tai
- 21__S-1__1Av1_1-audio1_1
 - -audio1-tunnus on käytössä, jos äänitallenne on merkitty videotallenteen
lisätallennintiedostoksi
 - Äänitallenteita numeroidessa otetaan huomioon vain äänitallentimien määrä: videokameroiden
tai niiden ääniraitojen määrää ei huomioida

Jos tapahtumasta on vain äänitallenteita, voidaan tapauskohtaisesti käyttää joko lisätallennintun-
nusta tai uutta mediatunnusnumeroa, esimerkiksi äänitallenteet

- 131__S-1__1X1_1 ja
- 131__S-1__1X1_1-audio2_1 vastaavat sisällöiltään toisiaan.

HUOM! Ensisijaiseksi äänitallentimeksi valittua tallenninta ei merkitä tunnuksella "-audio1".

Tallenteiden ryhmittely on aina tapauskohtaista johtuen esimerkiksi mahdollisista muistikortinvaih-
doista tai tallennuskatkoksista tapahtuman aikana. Mikäli tallennekokonaisuuksien hahmottami-
nen tai ryhmittäminen on hankalaa, on luontevaa käyttää uutta medianumeroa, esimerkiksi
tiedostojen

- 21__S-1__1Av1_1
- 21__S-1__1Av1_2a ja
- 21__S-1__1Av1_2b sisältöjä voi vastata videotiedostot
 - 21__S-1__1Av2_1
 - 21__S-1__1Av2_2
 - 21__S-1__1Av2_3a ja
 - 21__S-1__1Av2_3b sekä äänitiedostot
 - 21__S-1__1X1_1 ja
 - 21__S-1__1X2_1

Lisätallennintunnus ei ole käytössä, koska on annettu uudet mediatunnukset. Uuden mediatun-
nuksen antaminen on mahdollista vain uudessa arkistonmuodostustavassa.

Tiedostojen ”21__S-1__1Av1_1” ja ” 21__S-1__1Av2_1” ei tarvitse vastata toisiaan niin tarkasti, kuin lisätallennintunnusta käytettäessä: esimerkiksi vaihtoehtoiset kamerakulmat samasta tilanteesta, jotka eivät ajallisesti vastaa kovin tarkasti toisiaan.

HUOM! Lisätallennintunnusta voi seurata lisä- ja käyttökopiotunnus.

Lisätunnukset: liite-, litterointi ja indeksitiedostot

Lisätunnukset mahdollistavat aineistoa täydentävien tietojen tallentamisen alkuperäisten objektien yhteyteen. Tiedostonimessä voi olla korkeintaan yksi lisätunnus, jota voi edeltää lisätallennintunnus, ja jonka jälkeen voi olla käyttökopiotunnus. Lisä- ja käyttökopiotunnuksen esiintyminen peräkkäin on kuitenkin hyvin harvinaista, ellei jopa olematonta: Litterointi- ja indeksitiedostot ovat lähtökohtaisesti käyttökopioita. Ne on myös tallennettu usein monessa eri formaatissa säilymisen ja saavutettavuuden varmistamiseksi. Liitetiedostoissa moinen ketjutus on mahdollista, jos esimerkiksi ”_liite”-tunnuksella nimetyt valokuvat ovat sen verran laadukkaita, että käyttökopioiden tekeminen on järkevää.

Liitetiedosto on yleistermi kaikille tiedostoille, joilla on lisätunnus. Litterointi- ja indeksitiedostoille on määritetty omat tunnukset: ”-litt” ja ”-indx”. Näillä tiedostoilla on aina oma yksiselitteinen roolinsa suhteessa aineiston sisältöön, ja ne ovat mahdollisimman vaivattomasti saavutettavissa, kun ne liitetään alkuperäisen aineiston rinnalle useassa eri tiedostomuodossa.

Liitetunnuksella ”_liite” merkitty tiedosto puolestaan voi olla käytännössä mikä tahansa objekti, jota ei ole eriytetty omaksi arkistoyksiköksen, esimerkiksi kuva fyysisestä äänitemediasta, jonka sisältö on digitoitu, tai tallennetun tapahtuman ohjelmatietoliite. Tällaiset liitetiedostot sisältävät usein tietoa, jota ei aina ole mahdollista tai tarkoituksenmukaista yksiselitteisesti kirjoittaa kuvailevan metadatan järjestelmiin.

HUOM! Lisätunnusten numero-osio ei ala alaviivalla ja on siksi nimetty liitenumero-osioksi.

Liitetunnus

Liitetunnus muodostuu merkkijonosta ”_liite”, johon liitetään liitenumero-osio.

Liitetiedostot voivat liittyä mihin tahansa tiedostoon tai hakemistoon. Poistamalla tiedostonimestä liitetunnus saadaan digitaalinen objekti tai objektit tai aineistotaso, johon liite liittyy. Esimerkiksi tiedostot

- 1Vv1_1_liite1, 1Vv1_1_liite2-1 ja 1Vv1_1_liite2-2 liittyvät digitaaliseen objektiin 1Vv1_1
 - Ne eivät enää liity mahdollisiin muihin arkistoyksikön digitaalisiin objekteihin, kuten esimerkiksi 1Vv1_2, 1Vv1_3, 1Vv2_1, 1Vv3_1 jne.
 - Yksittäisten valokuvien tapauksissa liitetiedostot voivat olla esimerkiksi kuvan kääntöpuolen skannauksia, silloin kun niihin on kirjoitettu kuvaan liittyviä metatietoja
- 21Vv1_liite1 liittyy arkistoyksikön 21Vv1 objekteihin
 - Liitteet eivät enää liity sarjan mahdollisiin muihin arkistoyksiköihin ja niiden sisältämiin objekteihin kuten esimerkiksi 21Vv2_1, 21Vv3_1 jne.
- 321Md_liite1 liittyy arkiston 321 sarjatasoon Md
 - Tiedot/huomiot ovat yhteisiä kaikille sarjan objekteille 321Md1, 321Md2 jne.
- 4321__S-1__liite1 liittyy sarjatasoon S.1
 - Tiedot/huomiot ovat yhteisiä kaikille arkiston 4321 sarjan S.1 sisältäville objekteille
 - Huomaa sarjaerotinmerkin ”__” jälkeinen alaviiva ”_”, joka kuuluu liitetunnusosiioon

Litterointitunnus

Litterointitunnus muodostuu merkkijonosta ”-litt”, johon liitetään liitenumero-osio.

Litterointitiedostot sisältävät ääni- tai videotallenteen litteroinnin. Litterointitiedostot liittyvät aina tiettyyn haastatteluun tai tapahtumatallenteeseen: Yksi litterointitiedosto voi vastata yhtä tai useampaa haastattelu- tai tapahtumatallennetiedostoa riippuen siitä onko litteroitu kokonaisuus yhdessä vai useassa tiedostossa. On myös mahdollista, että esimerkiksi jokaiselle ääni- tai videotallenteelle on luotu omat litterointitiedostonsa.

Poistamalla tiedostonimestä litterointitunnus saadaan digitaalinen objekti tai objektit, johon litterointitiedosto liittyy, esimerkiksi tiedosto

- 1X1_1-litt1 sisältää kokonaisuuden 1X1_1 litteroinnin
 - Kokonaisuus 1X1_1 voi olla yksittäinen tiedosto tai esimerkiksi kahden tai useamman tiedoston kokonaisuus, esimerkiksi
 - 1X1_1a ja 1X1_1b tai
 - 1X1_1-1, 1X1_1-2, 1X1_1-3

Litterointitiedoston liittäminen sarjatasolle ei ole mielekäästä.

Litterointitiedostoissa on todennäköisesti aina useita saman nimisiä tiedostoja – vain tiedostomuoto ja sen päätte muuttuu. Esimerkiksi, jos alkuperäinen litterointi on tehty

- Pages-ohjelmalla (tiedostopäätte: .pages) tai
- Word-ohjelmalla (tiedostopäätte: .doc tai .docx)

on yllä olevien tiedostojen rinnalle lisäksi on arkistoitu samannimiset tiedostot myös teksti-, PDF- ja EPUB-muodoissa, ja eri tiedostojen nimet poikkeavat toisistaan vain tiedostopäätteiden osalta.

HUOM! Yllä olevissa esimerkeissä on arkistoitu yhteensä neljä eri tiedostoa, joilla on sama tietosisältö, ja joista vähintään yhden tulisi olla PAS-palvelun määritysten mukainen säilytys- tai siirtokelpoinen tiedostomuoto.

Indeksitunnus

Indeksitunnus muodostuu merkkijonosta ”-indx”, johon liitetään liitenumero-osio.

Indeksitiedostot sisältävät ääni- tai videotallenteen sisällön litterointia vähäisemmällä tarkkuudella. Indeksitiedostot liittyvät aina tiettyyn haastatteluun tai tapahtumatallenteeseen: Yksi indeksitiedosto vastaa yhtä tai useampaa haastattelu- tai tapahtumatallennetiedostoa riippuen siitä onko indeksoitu kokonaisuus yhdessä vai useassa tiedostossa. On myös mahdollista, että esimerkiksi jokaiselle ääni- tai videotallenteelle on tehty omat indeksitiedostonsa.

Poistamalla tiedostonimestä indeksitunnus saadaan digitaalinen objekti tai objektit, johon indeksitiedosto liittyy, esimerkiksi tiedosto

- 1X1_1-indx1 sisältää kokonaisuuden 1X1_1 indeksin
 - Kokonaisuus 1X1_1 voi olla yksittäinen tiedosto tai esimerkiksi kahden tai useamman tiedoston kokonaisuus, esimerkiksi
 - 1X1_1a ja 1X1_1b tai
 - 1X1_1-1, 1X1_1-2, 1X1_1-3

Indeksitiedoston liittäminen sarjatasolle ei ole mielekäästä: Sarjan tai muun vastaavan laajemman kokonaisuuden sisältötiedot kuuluvat kuvailevan metadatan järjestelmään. Tarvittaessa sama tai samankaltainen tieto voidaan lisätä sähköiseen arkistoon liitetiedostona käyttäen liitetunnusta.

Litterointitiedostoissa on todennäköisesti aina useita saman nimisiä tiedostoja – vain tiedostomuoto ja sen päätte muuttuu. Esimerkiksi, jos alkuperäinen indeksi on tehty

- Pages-ohjelmalla (tiedostopäätte: .pages) tai
- Word-ohjelmalla (tiedostopäätte: .doc tai .docx)

on yllä olevien tiedostojen rinnalle lisäksi on arkistoitu samannimiset tiedostot myös teksti-, PDF- ja EPUB-muodoissa, ja eri tiedostojen nimet poikkeavat toisistaan vain tiedostopäätteiden osalta.

HUOM! Yllä olevissa esimerkeissä on arkistoitu yhteensä neljä eri tiedostoa, joilla on sama tietosisältö, ja joista vähintään yhden tulisi olla PAS-palvelun määritysten mukainen säilytys- tai siirtokelpoinen tiedostomuoto.

Käyttökopiotunnus (__kk + numero tai __kk0 + lisätunnus)

Käyttökopiotunnus koostuu merkkijonosta ”__kk”, johon liitetään numero.

HUOM! Käyttökopiotunnuksen ”__kk0” jälkeen voidaan liittää lisätunnus.

__kk[1-9]

Käyttökopiotunnusta käytetään tilanteissa, joissa digitaalisesta arkistokappaleesta on tehty käyttökopio. Käyttökopiotunnuksen numero 1–9 ilmaisee tapauskohtaisesti määrittelemättömällä tarkkuudella, onko tiedostoa pakattu tai pienennetty vähän (1) vai paljon (9).

HUOM! Käyttökopioita ei lähtökohtaisesti talleteta PAS-palveluun, vaan ne ovat arkiston asiakaspalvelukäytössä. Poikkeuksena PAS- ja arkistointikäytänteisiin on tunnus ”__kk0”.

__kk0

Tunnusta ”__kk0” käytetään tilanteissa, joissa objekti tietoisesti digitoidaan niin sanotulla käyttökopiolaadulla. Tunnus ”__kk0” tarkoittaa, että aineistosta ei ole saatavana parempilaatuista digitaalista kopiota. Mutta toisaalta ”__kk0”:n puuttuminen tiedostonimestä ei takaa, että objekti olisi PAS-säilytyskelpoinen tai hyvälaatuinen. Arkistoon otetaan vastaan aineistoa suoraan asiakkailta. Tilanteesta riippuen tiedostoja ei aina kyetä nimeämään uudestaan aikaisemmin tässä dokumentissa määritettyjen sääntöjen mukaan arkistonmuodostuksen yhteydessä, eikä niiden PAS-säilytyskelpoisuutta pystytä aina arvioimaan. Näissä tilanteissa ei tiedostonimissä oteta kantaa tiedostojen laatuun eikä formaattiin. Eikä täten tiedostoihin merkitä mitään käyttökopio- tai muutakaan tunnusta: tiedostoja kohdellaan alkuperäisinä objekteina, ja niistä tapauskohtaisesti muodostetaan tai jätetään muodostamatta PAS-palvelun määrittämät säilytys- tai siirtokelpoiset tiedostot.

Poikkeuksena edelliseen on tilanteet, joissa asiakkaalta on tullut useita versioita samasta objektista. Tällöin arkistoinnin yhteydessä tapauskohtaisesti mahdollisesti merkitään käyttökopioiksi luokiteltavia tiedostoja tunnuksilla ”__kk1–9”. Tiedosto, jossa on paras laatu, jätetään aina merkitsemättä käyttökopiotunnuksella ”__kk0” tai ”__kk1–9”.

HUOM! Asiakkailta vastaanotettujen objektien nimistä kuitenkin aina poistetaan tai muunnetaan tulostumattomat sekä kielletyt US-ASCII-merkit osana nauhavarmennus- tai PAS-palveluun lähettämisen työnkulkua, ks. yllä "Musiikkiarkiston kokoelmat ja arkistot" ja alla "Syntysähköisen tai digitaalisen muodossa vastaanotetun aineiston erikoistapaukset".

Asiakkailta tulleista digitaalisista objekteista on myös mahdollista tehdä käyttökopiot – joita ei ole tarkoitus siirtää PAS-järjestelmään –, ja merkitä tapauskohtaisesti tiedostoihin käyttökopiotunnus ”__kk1–9”. Tunnusta ”__kk0” ei näissä tapauksissa käytetä. Tunnusta ”__kk0” käytetään vain arkiston omissa prosesseissa syntyneiden tiedostojen nimissä.

Tunnuksella ”__kk0” merkitty tiedosto voi olla tiedostomuodoltaan PAS-säilytyskelpoinen, mutta jo lähtökohtaisesti on tiedossa, että digitaalisen objektin tulisi olla huomattavasti laadukkaampi, jotta se olisi niin sanotusti arkistokelpoinen, esimerkiksi valokuvan suurempi pikselikooltaan tai bittisyvyydeltään tai äänitiedoston tarkempi näytteenottotaajuudeltaan tai suurempi bittisyvyydeltään.

Tunnuksen ”__kk0” perässä voi vielä olla lisätunnus. Muiden käyttökopiotunnusten perässä ei voi olla mitään muita tunnuksia: kyseiset tiedostot ovat aina muunnoksia muista arkistoiduista tiedostosta, jotka on nimetty sääntöjen mukaan, joten esimerkiksi lisä- tai osaobjektitunnusten on esiinnyttävä tiedostonimissä ennen käyttökopiotunnuksia.

HUOM! Määrittelemättä on osaobjektitunnuksista kootun käyttökopion nimeäminen. Onko esimerkiksi kolmesta tiedostosta, ”1Vk1_1a”, ”1Vk1_1b” ja ”1Vk1_1c”, yhdistetty käyttökopio nimeltään

- 1Vk1_1
- 1Vk1_1__kk0
- 1Vk1_1__kk1
- 1Vk1_1__kkk (käyttökopiokomposiitti)
- 1Vk1_1a-c
- 1Vk1_1a-c__kk0
- 1Vk1_1a-c__kk1
- 1Vk1_1a-c__kkk (käyttökopiokomposiitti)
- vai mikä?

Yllä mainitun kaltaisissa tilanteissa on luontevaa arkistoida käyttökopio arkistokopioiden yhteyteen, jolloin esimerkiksi useassa osassa digitoidut suuret julisteet ovat kokonaisuudessaan nähtävissä yhdessä kuvassa.

Sensuroitu käyttökopiotunnus (__skk[+ numero])

Tunnusta ”__skk” käytetään tilanteissa, joissa arkistokopiosta on sensuroitu osia esimerkiksi luovutussopimuksesta tai tietosuojalaista johtuen. Tunnuksen perässä ei ole numeroa, mikäli tiedosto

on johdettuarkistoon luovutetusta tai arkistossa digitoidusta arkistokopiosta alkuperäistä vastaavalla laadulla.

Pakatut käyttökopiot merkitään yllä olevan tunnuksen ”__kk” tavoin, esimerkiksi {pääosio}{pe-
rusosio}__skk3

Alkuperäiset tiedostot voidaan säilyttää päähakemistossa sensuroitujen käyttökopioiden rinnalla tai erillisessä hakemistossa: {sarjatunnus}---orig. Jälkimmäisessä tapauksessa mitään arkaluonteista tietoa ei saa sisältyä tiedostoihin, jotka sijaitsevat päähakemistossa, joten hakemiston voi antaa huoletta asiakaskäyttöön. Myöskään käyttökopiohakemistot eivät koskaan saa sisältää mitään arkaluonteisia tietoja.

HUOM! Tiedostojen alkuperäisyyttä ei ole merkitty tiedostonimiin. Jos metatietoihin, luovutusluetteloon tai liitetiedostoon ei ole merkitty mitä tiedostoja on otettu vastaan, tai missä järjestyksessä tiedostoja on luotu, ei pystytä olemaan varmoja mikä arkistoyksikön tiedostoista on alkuperäinen ja mikä muunnos, jos arkistossa on useita saman nimisiä eriformaattisia tiedostoja.

HUOM! Tapauksissa, joissa arkistoyksikössä on useita eriformaattisia tiedostoja, on todennäköistä, että alkuperäinen tiedostomuoto ei ole PAS-säilytyskelpoinen. Esimerkiksi PAS-siirtokelpoisia MP3-tiedostoja on saatettu muuttaa PAS-säilytyskelpoiseen WAV-muotoon arkistoinnin yhteydessä: on arkistoitu samannimiset tiedostot WAV- ja MP3-muodoissa. On myös mahdollista, että asiakkaalta on saatu sekä WAV- että MP3-muotoiset tiedostot, jotka molemmat on arkistoitu. Tilanteita, joissa on saatu useita eri formaattisia tiedostoja ei pystytä jälkikäteen erottamaan tilanteista, joissa formaattimuunnos on tehty arkistoinnin yhteydessä, jos toimenpiteistä ei ole kirjattu metatietoja, tai asiakkaan luovuttamien tiedostojen listausta ei ole tallennettu esimerkiksi liitetiedostona digitaalisten objektien yhteyteen.

Tähän päättyy tiedostonimiosuus. Tiedostonimi päättyy aina numeroon tai pieneen kirjaimen.

Syntysähköisen tai digitaalisessa muodossa vastaanotetun aineiston erikoistapaukset

Alkuperäisten hakemistorakenteiden sekä tiedostonimien säilyttäminen on suotavaa tilanteissa, joissa aineiston järjestäminen tai luetteloiminen olisi erittäin työlästä, tai alkuperäinen hakemistorakenne tiedostonimien avulla auttaa paremmin ymmärtämään aineistoa. Alkuperäisen aineiston indeksitiedoston voi liittää aineiston liitetiedostoksi, jolloin voidaan vielä erikseen tarkastella muuttamattomia alkuperäisiä tiedostonimiä tarpeen vaatiessa.

Alkuperäinen tiedostonimi säilytetään lisäämällä sen eteen sarjatunnus, arkistoyksikkönumero, mediatunnus ja alaviiva ”_”. Alaviiva vastaa raitaerotinta. Raitaerottimen jälkeen voidaan alkuperäiset tiedostonimet säilyttää siistityissä muodoissansa; nimissä käytetään vain sallittuja tulostuvia US-ASCII-merkkejä, ks. "Musiikkiarkiston kokoelmat ja arkistot". Siistitty alkuperäinen tiedostonimi vastaa raitatunnusta.

Alihakemistorakenne voidaan säilyttää lisäämällä merkkijono "---" edeltävien osien perään, minkä jälkeen jälleen siistitään seuraavat hakemisto- ja tiedostonimet pitkäaikaissäilytykseen soveltuvan muotoon. Täten voidaan säilyttää esimerkiksi useita "qwerty1.jpg"-nimisiä tiedostoja useissa eri hakemistoissa arkistosta tai sarjasta riippumatta, sillä alkuperäisten nimen eteen lisätään aina ainutlaatuiset arkistoyksikkötunnukset, mikä takaa digitaalisen objektin tiedostonimen ainutlaatuisuuden.

Nimeämisen syntaksi on muotoa: {Arkistoyksikkötunnus}_{alkup. nimi}---{alkup. nimi}].[tiedostopääte]]. Ensimmäisissä hakasulkeissa oleva alkuperäinen nimi -osio voi tarpeen vaatiessa toistua useita kertoja. Vähintään yksi raitaerottimen jälkeinen alkuperäinen nimi -osio on annettava; muutenhan aineisto olisi nimetty täysin uusiksi. Jälkimmäisissä hakasulkeissa olevaa tiedostopääteosiota ei hakemistonimissä käytetä. Tiedosto- tai hakemistonimi voi päättyä ensimmäisen tai n:n alkuperäinen nimi -osion jälkeen. Tiedostonimi päättyy, kuten tavallista, pistettä seuraavaan tiedostopäätteeseen. Esimerkiksi hakemisto "434__H-3__1Z1_APURAHA_ASIAT" sisältää tiedostot

- 434__H-3__1Z1_APURAHA_ASIAT--001p.jpg

- 434__H-3__1Z1_APURAHA_ASIAT---apuraha_selvitykseen__2009.doc
- 434__H-3__1Z1_APURAHA_ASIAT---arnold.jpg

Tässä tapauksessa alkuperäinen hakemistonimi on ollut "APURAHA_ASIAT", jonka eteen on lisätty arkistotunnus "434", sarjatunnus "H-3" ja arkistoyksikkötunnus "1Z1" erotinmerkkeineen "__" ja "_". Tämän jälkeen on alkuperäiset hakemisto- ja tiedostonimet voitu säilyttää siisityissä muodoissansa erontinmerkkijonon "---" jälkeen.

Mediatunnuksena voidaan käyttää homogeenisen aineiston tapauksissa aineistoa vastaavaa mediatunnusta. Sekalaisen aineiston tapauksissa on suotavaa käyttää mediatunnusta "Smj" tai "Z".

HUOM! On olemassa myös toinen tapa, jossa arkistoyksikköosio on jätetty pois, esimerkiksi hakemisto "433__S-6-1__FOLDER01" sisältää tiedostot

- 433__S-6-1__FOLDER01---ZOOM0001.WAV
- 433__S-6-1__FOLDER01---ZOOM0002.WAV

Arkistoyksikkö- ja mediatunnuksen pois jättäminen sallitaan tapauksissa, joissa objektien nimet ja hakemistorakenne ovat hyvin selkeitä ja ymmärrettäviä.

HUOM! Syntysähköinen tai sähköisesti vastaanotettu aineisto voi olla hyvin monimuotoista ja sekavasti järjestettyä. Esimerkiksi Dicientian arkiston (tunnus: 361) siirretyt optiset levyt sisältävät noin 200 000 tiedostoa, ja levyjen hakemistorakenne on joissain tapauksissa hyvin sekava. Lisäksi hakemisto- ja tiedostonimet saattavat sisältää useita pisteitä tai tiedostonimet eivät sisällä yhtäkään pistettä: tiedostopäätettä ei ole. Näissä tapauksissa aineisto järjestetään ja objektien nimet käsitellään vallitsevien sääntöjen mukaan, mutta esimerkiksi PAS-palveluun lähetettäessä tunnistamattomien sekä PAS-palvelun tukemattomien objektien SIP-paketointi on toistaiseksi mahdotonta (20220513).

Erikoistunnukset

Erikoistunnukset merkitään tiedostonimen loppuun tiedostopäätteen perään.

Vauriotunnus: ---dmgd

Vaurioituneen tiedoston nimen loppuun lisätään merkkijono "----dmgd".

Aineisto voi olla vaurioitunutta. Esimerkiksi ZIP-paketit on mahdollista purkaa vain osin tai ei lainkaan. Tällöin alkuperäisen ZIP-tiedoston nimen perään lisätään vaurion ilmaiseva tunnus "dmgd" erotinmerkkijonon "---" jälkeen, esimerkiksi "__Y-E-Ker__4__1Smj1_arkisto.zip---dmgd".

Vauriotunnus "dmgd" ilmaisee myös tuntemattoman tai tukemattoman paketin. Sitä käyttäen merkitään myös paketit, joita ei saada lainkaan auki, mutta on syytä olettaa, että kyseessä on jollain työkalulla purettava paketti.

Objektit, jotka on mahdollista purkaa paketista, otetaan talteen ja nimetään vallitsevien käytänteiden mukaisesti.

Vauriotunnusta on mahdollista käyttää myös digitoitujen aineistojen yhteydessä, mutta tällöin yhden tai useamman seuraavista ehdoista tulee täytyttyä:

- Yhtään eheää digitaalista kopioita objektista ei ole olemassa
 - Aineiston varmuuskopiointi on epäonnistunut
- Alkuperäisen analogisen aineiston digitointi on mahdotonta tai erityisen hankalaa
 - Alkuperäinen aineisto on esimerkiksi palautettu omistajalle tai hävitetty
- Korvaavan digitoinnin täydellisyyttä tai laatua ei voida luotettavasti taata
 - Alkuperäinen aineisto on vahingoittunut, ja on mahdollista, että olemassa olevassa objektissa on ainakin ajoittain arkisto- tai käyttökelpoisia kohtia

Migrointitunnus: ---migrated-<vvvkkpp>[-osaobjektitunnus]

Migroidun korvattavan tiedoston nimen loppuun ennen tiedostopäätettä lisätään merkkijono "---migrated-<vvvkkpp>[-osaobjektitunnus]".

Korvattavan objektin status on päivitettävä Skaalaan PAS-kirjanpitoon. Uusi objekti saa korvattavan tiedoston tiedostonimen, ja sille on luotava omat pysyvät UUID- ja UUID-SHORT-tunnisteensa

Musiikkiarkiston PAS-käytänteiden mukaisesti, ks. dokumentti "Aineiston hallinta Musiikkiarkistossa".

HUOM! Migroitu tiedosto on edelleen aktiivisessa säilytyksessä niin kauan kunnen se poistetaan. Vain viimeisimmän migraatiotapahtuman päivämäärä jää muistiin objektin nimeen; mahdolliset aikaisemmin suoritettut migraatiotapahtumat ovat muiden suoritettujen PREMIS-tapahtumien tavoin tallessa Skaalan PAS-kirjankirjanpidossa.

Korvatun tiedoston nimi, voi olla esimerkiksi "1X1_1---migrated-20200101.wav", mikä tarkoittaa, että tiedostomigraatio on suoritettu 1.1.2020. Korvaava objekti saa korvattavan objektin nimen: "1X1_1.wav".

HUOM! Mikäli migroidaan kertaalleen korvattu objekti "1X1_1---migrated-20200101.wav" esimerkiksi 31.12.2020, syntyy aikaisemman migraation tuloksena luodun objektin "1X1_1.wav" rinnalle, uusi objekti, joka merkitään migraatio-osaobjektitunnuksella "-a". Huomaa migraatio-osaobjektitunnusta edeltävä yhdysviiva "-". Aktiivisessa säilytyksessä on tällöin kolme objektia:

- 1X1_1.wav, joka on ensimmäisen migraation korvaava objekti
 - Objekti 1X1_1---migrated-20200101.wav on saanut uuden nimen 1X1_1---migrated-20203131.wav
- 1X1_1---migrated-20203131.wav, jonka korvaava vastinpari on objekti 1X1_1-a.wav
 - Objektit 1X1_1.wav ja 1X1_1-a.wav on migroitu samasta tiedostosta

Vrt. Osaobjektitunnus yllä, arkistossa 445 olevat vaihtoehtoiset digitointilaadut "445__H-3__1Y1" ja "445__H-3__1Y2": vain osaobjektitunnus "a" on käytössä.

HUOM! Mikäli migroidaan sekä kertaalleen korvattu objekti "1X1_1---migrated-20200101.wav" että aikaisemman migraation korvaava objekti "1X1_1.wav" esimerkiksi 31.12.2020, syntyy kaksi uutta objektia, jotka merkitään osaobjektitunnuksilla "-a" ja "-b". Myös migroitavat korvattavat objekti merkitään osaobjektitunnuksilla "-a" ja "-b", koska objektien nimeen liitetään migraatiotapahtuman päivämäärä. Aktiivisessa säilytyksessä on tällöin neljä objektia

- 1X1_1---migrated-20203131-a.wav, jonka korvaava vastinpari on objekti 1X1_1-a.wav
 - Korvattavan objektin alkuperäinen nimi voi olla 1X1_1.wav tai 1X1_1---migrated-20200101.wav
- 1X1_1---migrated-20203131-b.wav, jonka korvaava vastinpari on objekti 1X1_1-b.wav
 - Korvattavan objektin alkuperäinen nimi voi olla 1X1_1.wav tai 1X1_1---migrated-20200101.wav

HUOM! Korvattavan objektin migraatio-osaobjektitunnuksen edellä ei ole välttämätöntä käyttää yhdysviivaa, sillä samanlaista sekaantumisvaaraa osaobjektitunnuksen kanssa ei ole, sillä mahdollinen osaobjektitunnus olisi ennen merkkijonoa "---migrated-20203131-a.wav" ja "---migrated-20203131-b.wav" heti numero-osion perässä.

Poistetun objektin tunnus: ---deleted-<vvvkkpp>

Poistetun tiedoston nimen loppuun lisätään merkkijono "---deleted-<vvvkkpp>".

Objektin status on päivitettävä Skaalaan PAS-kirjanpitoon, ks. dokumentti "Aineiston hallinta Musiikkiarkistossa".

Poistetun tiedoston nimi, voi olla esimerkiksi "1X1_1.wav---deleted-20200101", mikä tarkoittaa, että tiedoston poistaminen on suoritettu 1.1.2020.

Liitteet

Liite 1. Tiedostojen nimeämisesimerkkejä

Vanha arkistonmuodostustapa

Tiedostonimen osiot

- pääosio
 - arkistotunnus
 - vain, jos kyseessä on henkilö- tai yhteisöarkisto
 - mediatunnus
- perusosio
 - arkistoyksikkötunnus
 - osakokonaisuustunnus
- tiedosto-osio
 - numero-osio
 - osaobjektitunnus
 - migraatio-osaobjektitunnus
 - lisätallennintunnus
 - liite-, litterointi tai indeksitunnus
 - käyttökopiotunnus
 - vaurio-, migraatio- tai poistetun objektin tunnus (erikoistunnukset)
 - Poistettujen objektien tunnukset ovat mahdollisesti tallessa vain Skaalan PAS-kirjanpidossa tai kuvailevan metatadan järjestelmissä

Yleiskokoelmien mediakokoelmat

Mediatunnus + juokseva numero + numero-osio (pakolliset vähimmäisvaatimukset)

- Av1_1
- Av2_1, Av2_2
- Ck23_1-1, Ck23_1-2

Mediatunnus + juokseva numero + osakokonaisuustunnus + numero-osio

- Ck74A_1, Ck74A_2
- Ck74B_1
- Ck74C_1

Mediatunnus + juokseva numero + numero-osio + osaobjektitunnus

- X123_1a, X123_1b

Mediatunnus + juokseva numero + osakokonaisuustunnus + numero-osio + osaobjektitunnus

- 313Ck1A_1a, 313Ck1A_1b
- 313Ck1B_1a, 313Ck1B_1b, 313Ck1B_1c

Esimerkkejä hakemistossa Av (sarjassa) mahdollisesti olevista hakemistoista ja tiedostoista.

- Hakemistossa Av1 (Av-sarjan ensimmäinen arkistoyksikkö) olevat tiedostot
 - Av1_1
 - Av1_2a, Av1_2b
 - Av1_3
- Hakemistossa Av2A (Av-sarjan toisen arkistoyksikön ensimmäinen osa) olevat tiedostot
 - Av2A_1
 - Av2A_1_liite1
- Hakemistossa Av2B (Av-sarjan toisen arkistoyksikön toinen osa) olevat tiedostot
 - Av2B_1a
 - Av2B_1b

Lisäksi on mahdollista, että sarjahakemistossa ”Av” on ”Av2”-kokonaisuudelle yhteisiä liitetiedostoja, esimerkiksi ”Av2_liite1”. Useille eri hakemistoihin sijoitetuille kokonaisuuksille yhteisiä liitetiedostoja ei saa sijoittaa osakokonaisuushakemistoihin, koska tällöin tiedostonimen alkuosa ei täsmäisi hakemistonimen kanssa.

Kaikkien edellisten esimerkkien perään voidaan vielä liittää lisätunnus, lisätallennintunnus tai käyttökopiotunnus, esimerkiksi

- tiedostoon Av1_1 voi liittyä tiedostot
 - Av1_1_liite1, Av1_1_liite2-1, Av1_1_liite2-2
 - Av1_1-indx1, Av1_1-indx2
 - Av1_1-litt1
 - Yhteen objektiin voi liittyä useita liite-, litterointi- sekä indeksitiedostoja
- tiedostojen Av2_1 ja Av2_2 yhteyteen tallennettu tiedosto
 - Av2_1_liite1 liittyy vain tiedostoon Av2_1
 - Av2_liite1 liittyy kaikkiin arkistoyksikön Av2 tiedostoihin
 - *HUOM! Tiedostonimessä on vain yksi alaviiva, joka on osa liitetunnusta. Mediatunnuksen jälkeinen juokseva numero ei sisällä alaviivaa*
- tiedostojen Av23_1, Av23_2 yhteyteen tallennettu tiedosto
 - Av23-audio1_1 sisältää molemmille tiedostoille yhteisen lisä-äänitallenteen
 - Av23_1-audio1_1 sisältää vain tiedoston Av23_1 lisä-äänitallenteen
 - Arkistoyksikön numero-osio ”_1” erottaa lisä-äänitallenteet toisistaan, joten ”-audio1”-tunnusta on mahdollista käyttää molemmissa lisä-äänitallenteiden tiedostonimissä
 - Hahmottamisen helpottamiseksi on mahdollista käyttää toisessa tallenteessa ”-audio2”-tunnusta. Varsinkin, jos kyseessä on kahdella eri tallennusvälineellä äänitetyt tiedostot, jolloin eri tallennusvälineet ovat tunnistettavissa
- tiedostoista Av131_1a ja Av131_1b on tehty asiakaskäyttöön käyttökopiot
 - Av131_1a__kk3 ja Av131_1b__kk3

Henkilö- ja yhteisöarkistot

Tiedostojen nimeäminen vastaa lähes täysin yleiskokoelmien mediatunnussarjojen nimeämistä. Ainoana erona on henkilö-/yhteisöarkistotunnus, joka aloittaa tiedostonimen.

Arkistotunnus + mediatunnus + juokseva numero + numero-osio (pakolliset vähimmäisvaatimukset)

- 1Av1_1
- 1Av2_1, 1Av2_2

Pakollisten vähimmäisvaatimusten lisäksi tiedostonimet voivat sisältää osakokonaisuus-, osaobjekti-, lisä-, lisätallennin-, käyttökopio- tai erikoistunnuksia yllä esitettyjen yleiskokoelmien media-kokoelmien esimerkkejä mukaillen, esimerkiksi

- tiedostoon Av1_1 voi liittyä tiedostot
 - 1Av1_1_liite1, 1Av1_1_liite2-1, 1Av1_1_liite2-2
 - 1Av1_1-indx1, 1Av1_1-indx2
 - 1Av1_1-litt1
- tiedostojen 1Av2_1 ja 1Av2_2 yhteyteen tallennettu tiedosto
 - 1Av2_1_liite1 liittyy vain tiedostoon 1Av2_1
 - 1Av2_liite1 liittyy kaikkiin arkistoyksikön 1Av2 tiedostoihin

Uusi arkistonmuodostustapa

Tiedostonimen osiot

- pääosio
 - kokoelma- tai arkistotunnus
 - sarjatunnuserotin + sarjatunnus + sarjatunnuserotin
- perusosio
 - arkistoyksikkötunnus
 - osakokonaisuustunnus
- tiedosto-osio

- numero-osio
- osaobjektitunnus
- migraatio-osaobjektitunnus
- lisätallennintunnus
- liite-, litterointi tai indeksitunnus
- käyttökopiotunnus
- vaurio-, migrointi- tai poistetun objektin tunnus (erikoistunnukset)
 - Poistettujen objektien tunnukset ovat mahdollisesti tallessa vain Skaalan PAS-kirjanpidossa tai kuvailevan metatadan järjestelmissä

Yleiskokoelmien erikoiskokoelmat

Erikoiskokoelmatunnus + sarjatunnuserotin + sarjatunnus + sarjatunnuserotin + juokseva numero + mediattunnus + juokseva numero + numero-osio (pakolliset vähimmäisvaatimukset)

- Smj__1__1X1_1
 - Sisällön mukaan järjestetyn aineiston ensimmäisen sarjan ensimmäisen arkistoyksikön ensimmäinen tiedosto
- Tdt__1-1__1Y1_1 ja Tdt__1-1__1Y1_2
 - Levy-yhtiötiedotekokoelman ensimmäisen (pää)sarjan ensimmäisen (ala)sarjan ensimmäisen arkistoyksikön kaksi tiedostoa

Eri levy-yhtiöiden tiedotteet on järjestetty pääsarjoittain: sarja "Tdt__1__" sisältää eri yhtiön tiedotteet kuin sarja "Tdt__2__"

Pakollisten vähimmäisvaatimusten lisäksi tiedostonimet voivat sisältää osakokonaisuus-, osaobjekti-, lisä-, lisätallennin-, käyttökopio- tai erikoistunnuksia yllä esitettyjen yleiskokoelmien media-kokoelmien esimerkkejä mukailten. Esimerkiksi arkistoyksikkö "Smj__1__1" voi sisältää tiedostot

- Smj__1__1X1_1a, Smj__1__1X1_1b
- Smj__1__1X2_1, Smj__1__1X2_2
- Smj__1__1X3A_1, Smj__1__1X3B_1
- Smj__1__1Av1_1
 - Smj__1__1Av1_1_liite1
 - Smj__1__1Av1_1-indx1
- Smj__1__1Av1_1-cam2_1-1, Smj__1__1Av1__1-cam2_1-2
 - Smj__1__1Av1_1-cam2_1-indx1

Henkilö- ja yhteisöarkistot

Tiedostojen nimeäminen vastaa lähes täysin yleiskokoelmien erikoiskokoelmien nimeämistapaa. Ainoana erona on henkilö- tai yhteisöarkistotunnus, joka aloittaa tiedostonimen ja korvaa täten erikoissarjatunnuksen.

arkistotunnus + sarjatunnuserotin + sarjatunnus + sarjatunnuserotin + juokseva numero + mediattunnus + juokseva numero + numero-osio (pakolliset vähimmäisvaatimukset)

- 1__1__1X1_1
 - Arkiston numero 1 ensimmäisen sarjan ensimmäisen arkistoyksikön ensimmäinen tiedosto
- 2__1-1__1Y1_1 ja 2__1-1__1Y1_2
 - Arkiston numero 2 ensimmäisen (pää)sarjan ensimmäisen (ala)sarjan ensimmäisen arkistoyksikön kaksi tiedostoa

Pakollisten vähimmäisvaatimusten lisäksi tiedostonimet voivat sisältää osakokonaisuus-, osaobjekti-, lisä-, lisätallennin-, käyttökopio- tai erikoistunnuksia aikaisemmin esitettyjä esimerkkejä mukailten.

Liite 2. hakemisto- ja tiedostonimien osioiden säännölliset lausekkeet

HUOM! Ei testattu! (20200227)

Henkilö- ja yhteisöarkistotunnus

RegEx: (\d+)

Yleiskokoelman erikoiskokoelmatunnus

RegEx: (\D{1,5})

Sarjatunnus uudessa arkistonmuodostustavassa

RegEx: ([A-Za-z0-9-]+)

Sarjatunnuserotin ”__” -> RegEx: __([A-Za-z0-9_-]+)__

Sarjatunnus vanhassa arkistonmuodostustavassa (mediatunnus)

RegEx: (\D{1,5})

Arkistoyksikkötunnus vanhassa arkistonmuodostustavassa (juokseva numero)

RegEx: (\d+)

Arkistoyksikkötunnus uudessa arkistonmuodostustavassa (juokseva numero + mediatunnus + juokseva numero)

RegEx: (\d+\D{1,5}\d+)

Osakokonaisuustunnus

RegEx: ([A-Z]{1,2})?

HUOM! Käytännössä yksi kirjain, mutta poikkeuksena ainakin 34Kn, jossa kaksi kirjainta ei merkitse osakokonaisuustunnusta vaan jotain muuta.

Numero-osio

RegEx: (_\d_-)+

Osaobjektitunnus

RegEx: ([a-z])?

Migraatio-osaobjektitunnus

RegEx: (-[a-z])?

Lisätallennintunnus

RegEx: (-cam|-audio)([2-9]+\d_-)+?

Lisätunnukset

Liitetunnus

RegEx: (_liite\d_-)+?

Litterointitunnus

RegEx: (-litt\d_-)+?

Indeksitunnus

RegEx: (-indx[\d_-]+)?

Liitenumero-osio

RegEx: ([\d_-]+)

käyttökopiotunnukset

__kk0

RegEx: (__kk0[\w]+)?

__kk1-9

RegEx: (__kk[1-9]{1})?

Erikoistunnukset

Vauriotunnus: ---dmgd

Merkkijono: ---dmdg

Migrintitunnus: ---migrated-<vvvkkpp>**[-osaobjektitunnus]**

RegEx: ---migrated-(\d{8})[-a-z]

Poistetun objektin tunnus: ---deleted-<vvvkkpp>****

RegEx: ---deleted-(\d{8})

Liite 3. Säännölliset lausekkeet eri kokoelmien ja arkistojen tiedostonimille

Lienee järkevää tehdä vain muutama yleispätevä säännöllinen lauseke validoimaan tiedostonimiä uudessa ja vanhassa arkistonmuodostustavassa. Onko tarpeellista, sillä sääntöihin ja aineiston järjestämiseen liittyy paljon asioita, joita ei voi säännöllisillä lausekkeilla tai muutenkaan helposti tarkistaa, esimerkiksi sijaitsevatko tiedostot oikeiden kokoelmien alihakemistoissa ja vastaako arkistoon talletetun ja luetteloidun tunnus tiedostonimeä.