

# PAS-PALVELUIDEN RAJAPINNAT

v3.0.0

WHITE PAPER

Tämä määrittely on osa opetus- ja kulttuuriministeriön  
Avoimen tieteen ja digitaalisen kulttuuriperinnön kokonaisuutta



## **Lisenssi**

*Creative Commons Suomi CC-BY-NC-SA 4.0 (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.fi>)*

*Tämän määrittelyn käyttäjälle annetaan oikeus jakaa eli kopioida, levittää, näyttää ja esittää teosta sekä valmistaa muutettuja teoksia seuraavilla ehtoilla:*

- *Opetus- ja kulttuuriministeriö nimetään teoksen tekijäksi (ei kuitenkaan siten, että ilmoitus viittaisi lisenssinantajan tukevan lisenssinsaaajaa tai teoksen käyttötapaa).*
- *Epäkaupallinen osapuoli eli käyttäjä ei saa käyttää teosta kaupallisesti.*
- *Mikäli käyttäjä tekee muutoksia tai käyttää teosta omien teostensa pohjana, tulee johdannaisteos jakaa samalla tai samankaltaisella lisenssillä.*

## SISÄLTÖ

<b>1</b>	<b>JOHDANTO</b> .....	<b>5</b>
1.1	PAS-palvelut .....	5
1.2	Rajapinnat PAS-palveluissa .....	5
1.3	Rajapintojen automaatio .....	6
<b>2</b>	<b>RAJAPINNAT</b> .....	<b>7</b>
2.1	Autentikointi ja valvonta .....	7
2.2	Viestien rakenne .....	7
2.2.1	Kyselyviestit .....	7
2.2.2	Palautusviestit .....	8
2.3	Rajapinnan toiminnot .....	8
2.4	Estetyt tasot .....	9
<b>3</b>	<b>AINEISTON SIIRTO</b> .....	<b>10</b>
3.1.1	Lähetysten alustaminen .....	10
3.1.2	Tiedoston lähettäminen .....	10
3.1.3	Lähetysten tilan kysely .....	11
3.1.4	Siirtojen luettelointi .....	11
3.1.5	Vastaanoton tilan esittäminen .....	11
3.1.6	Siirtopaketin tarkastusraportin nouto .....	12
<b>4</b>	<b>AINEISTON HAKU</b> .....	<b>13</b>
4.1	Hakuviestit .....	13
4.2	Hakutoiminto .....	13
4.2.1	Hakutoiminnon ominaisuudet ja syntaksi .....	14
4.2.2	Hakulauseet .....	14
<b>5</b>	<b>SÄILYTETTÄVÄN AINEISTON HALLINTA</b> .....	<b>15</b>
5.1	Säilytyspaketin tiedot .....	15
5.2	Aineiston muodostus jakelupaketiksi .....	15
<b>6</b>	<b>AINEISTON JAKELU</b> .....	<b>17</b>
6.1	Muodostettavien ja noudettavien jakelupakettien listaus .....	17
6.2	Jakelupaketin seuranta ja niihin kohdistuvat komennot .....	17
6.3	Muodostettujen jakelupakettien noutaminen .....	17
<b>7</b>	<b>KÄSITTELYHISTORIA</b> .....	<b>18</b>
7.1	Siirtohistoria .....	18
7.2	Säilytyshistoria .....	18
7.3	Jakeluhistoria .....	18
<b>8</b>	<b>TILASTOT</b> .....	<b>19</b>
8.1	Yhteenveto tilastoista .....	19
<b>9</b>	<b>JULKISET AVAIMET</b> .....	<b>20</b>
9.1	PAS-palvelun jakeluavaimen noutaminen .....	20
<b>LIITE A.</b>	<b>ESIMERKKEJÄ PROSESSIN AUTOMATISOINNISTA</b> .....	<b>21</b>
	Automaatioprosessi aineistojen siirtämiseksi PAS-palveluun .....	21
	Automaatioprosessi aineistojen noutamiseksi PAS-palvelusta .....	21
<b>LIITE B.</b>	<b>SIIRTOPAKETIN TARKASTUSRAPORTTI</b> .....	<b>23</b>
	Objektit .....	23
	Tapahtumat .....	25

Agentit.....27

**LÄHTEET .....28**

## 1 JOHDANTO

Tämä dokumentti esittelee PAS-palveluiden rajapinnat, joiden avulla kansallisia PAS-palveluita hyödyntävät organisaatiot voivat mm. siirtää aineistoa säilytykseen PAS-palveluihin tai hakea niistä aineistoa. Rajapinnat on tarkoitettu käytettäväksi erityisesti automaattisesti toimivassa tietojärjestelmien välisessä integraatiossa. Rajapintojen suunnittelussa on tavoiteltu mahdollisimman yksinkertaista, mutta skaalautuvaa ja myöhemmin uusilla toiminnallisuuksilla täydennettävissä olevaa mallia.

Dokumentti on laadittu osana PAS-palveluiden rajapintojen uudistustyötä, joiden tarkoituksena on tuottaa rajapinnat-määrittelyistä uusi pääversio. Dokumentissa esitetään rajapintojen toiminnot ja kutsut, mutta se ei määrittele pyyntöjen syntaksia tai palautusviestejä tarkasti. Se ei siten ole määrittely, vaan suunnitteludokumentti.

### 1.1 PAS-palvelut

PAS-palveluilla tarkoitetaan kulttuuriperintöaineistojen ja tutkimusaineistojen pitkäaikaissäilyttämiseen tuotettuja palveluita yhdessä. Tässä dokumentissa hyödyntävällä organisaatiolla tarkoitetaan PAS-palvelua digitaalisten aineistojen pitkäaikaissäilyttämiseen käytettävää organisaatiota, laitosta tai muuta toimijaa.

Pitkäaikaissäilytyksellä tarkoitetaan digitaalisen informaation säilyttämistä ymmärrettävänä ja käyttökelpoisena useiden kymmenien ja jopa satojen vuosien ajan. Laitteet, ohjelmistot ja tiedostomuodot vanhenevat ajan myötä, mutta informaation täytyy säilyä. Luotettava pitkäaikaissäilyttäminen edellyttää sisällön eheyden aktiivista valvontaa ja monenlaisiin riskeihin varautumista. Tässä ovat keskeisessä asemassa metatiedot, jotka kuvailevat mm. aineiston sisällön, historian ja alkuperän sekä tiedot siitä, miten informaatiota voidaan käyttää.

Kulttuuriperintöaineistojen pitkäaikaissäilyttämiseen tuotettu PAS-palvelu (Kulttuuriperintö-PAS-palvelu) takaa kirjastojen, arkistojen ja museoiden keskeisten kansallisten digitalisten tietovarantojen pitkäaikaissäilyttämisen. Digitaalisilla kulttuuriperintöaineistoilla tarkoitetaan sekä digitoituja että digitaaliseen muotoon tuotettuja kulttuuriperintöaineistoja: lakisäateisen säilyttämisen piiriin kuuluvia digitaalisia kulttuuriaineistoja, kansalliseen kulttuuriperintöön kuuluvaa digitaalista asiakirjallista aineistoa sekä aineellisen ja henkisen kulttuuriperinnön säilyttämisestä vastaavien, opetus- ja kulttuuriministeriön hallinnonalalla toimivien organisaatioiden muita pitkäaikaissäilytyksen piiriin kuuluvia digitaalisia tietovarantoja.

Tutkimusaineistojen pitkäaikaissäilyttämiseen tuotettu PAS-palvelu (Fairdata PAS-palvelu) varmistaa tutkimuksen digitaalisten aineistojen saatavuuden ja pitkäaikaisen säilyvyyden. Tämä PAS-palvelu tukee osaltaan pysyvää ja koordinoitua toimintamallia tutkimusaineistojen hallinnan tueksi. Pyrkimyksenä on, että tutkimuksen todennettavuus ja toistettavuus elinkaaren eri vaiheissa onnistuu ja tulosten hyödyntäminen on helppoa. Tällöin tutkimustuloksia voidaan käyttää yhä uudelleen, arvioida, hyödyntää päätöksenteossa ja turvata digitalisoitumisen myötä yhä nopeammin kasvavat tietomäärät tulevien tutkijasukupolvien käyttöön.

### 1.2 Rajapinnat PAS-palveluissa

Siirrettävä aineisto paketoidaan *Aineistojen ja niiden metatietojen paketointi pitkäaikaissäilytykseen* [FI\_META] sekä *Säilytys- ja siirtokelpoiset tiedostomuodot* [FI\_FORMATS] -määrittelyiden mukaisesti.

Aineiston siirtäminen, hakeminen, muodostus jakelupaketeiksi ja noutaminen tapahtuu hyödyntäen HTTP-kutsuja [RFC\_7230] REST-käytäntöjen mukaisesti. Aineiston siirrosta syntyy aina HTML- ja PREMIS-muotoiset [PREMIS] tarkastusraportit. Tässä dokumentissa määritellään tähän viestintään tarvittavat HTTP-kysely- ja .vastausviestit [RFC\_7231]. Rajapinnan vastausviestien rungon rakenteena käytetään RFC 7159 -määrittelyn [RFC\_7159] mukaista JSON-muotoa lukuun ottamatta varsinaista säilytyksestä haettavaa aineistoa tai sen tarkastusraporttia. Jakelupaketit toimitetaan ZIP- tai TAR-pakettina, jakelupaketin metatiedot XML-muodossa, ja siirtopaketin tarkastusraportit XML- tai HTML-muodossa.

Virhetilanteissa tai muutoin poikkeavassa viestittelyssä käytetään HTTP-statuskoodistoa RFC 7231 -määrittelyssä [RFC\_7231] esitetyllä tavalla, mutta kuitenkin yleisesti vakiintuneet tavat ja käytännöt huomioiden, mikäli tässä määrittelyssä ei muuta määritellä.

Tässä dokumentissa hyödyntävän organisaation ohjelmistolla tarkoitetaan ohjelmistoa, joka kommunikoi PAS-palvelun kanssa käyttäen sovittuja tietoliikennetarkoituksia. Hyödyntävän organisaation ohjelmisto voi olla esimerkiksi organisaation kokoelmanhallintajärjestelmä tai jokin muu ohjelmisto.

Käyttäjätunnuksella tarkoitetaan tunnusta, joka identifioi hyödyntävän organisaation yksikäsitteisesti. Rajapinnan käyttö on aina rajattu käyttäjätunnuskohtaisesti niin, että hyödyntävällä organisaatiolla on tehtyjen sopimusten mukaisesti oltava tarvittavat oikeudet aineiston käsittelyyn.

### 1.3 Rajapintojen automaatio

PAS-palveluiden rajapintamäärittely on tehty siitä lähtökohdasta, että rajapintoja hyödynnetään täysin automaattisesti. Jos hyödyntävän organisaation tarkoituksena on esimerkiksi lähettää siirtopaketteja tai noutaa jakelupaketteja useita kappaleita (esimerkiksi kymmeniä tuhansia tai enemmän), voi hyödyntävän organisaation ohjelmisto automaattisesti luoda tarvittavat kyselyt, lähettää ne PAS-palvelulle ja käsitellä saamansa vastaukset. Jos hyödyntävällä organisaatiolla on tarpeen tehdä säännöllisin väliajoin (esimerkiksi päivittäin) tällaisia prosesseja rajapinnan avulla, tarkoituksena voi käyttää erilaisia ratkaisuja, esimerkiksi ajastustoimintoa, joka käynnistää prosessin hyödyntävän organisaation ohjelmistossa automaattisesti ja säännöllisesti. Esimerkkejä automaatioprosesseista on kuvattu liitteessä A. Yksityiskohtat rajapinnan toiminnasta on esitetty luvuissa [Error! Reference source not found.](#) ja [Error! Reference source not found.](#)

Deleted: 2

Deleted: 3

## 2 RAJAPINNAT

PAS-palvelujen ohjelmointirajapinnassa käytetään HTTP-kutsuja RESTin mukaisesti. Automaatiologiikan toteuttamisen helpottamiseksi, erityisesti koskien aineistojen hakua ja noutamista, PAS-palvelu tarjoaa palautusviesteissä suoraan ne komennot, joita seuraavassa vaiheessa tavallisesti tarvitaan. Tällöin riittää, että hyödyntävän organisaation ohjelmisto poimii saaduista palautusviesteistä nämä komennot ja välittää ne muuttamattomina takaisin PAS-palvelulle kyselyviesteinä. Tässä luvussa esitellyissä HTTP-kutsuissa käytetään seuraavia määreitä:

- `<base>` – Merkkijono: `https://pas.csc.fi/api/2.0`
- `<contract>` – Sopimustunniste, joka rajaa käsittelyn ainoastaan tietyn sopimuksen alaiseen aineistoon.
- Komennoilla voi olla myös muita määreitä, jotka on erikseen esitelty ko. viestin määrittelyn yhteydessä.

### 2.1 Autentikointi ja valvonta

Autentikoinnissa käytetään TSL/SSL-suojausta ja HTTP Basic Access -autentikointia. Hyödyntävän organisaation ohjelmisto autentikoidaan, minkä jälkeen se voi käyttää HTTP-kutsuissa hyödyntävälle organisaatiolle valtuutettuja sopimustunnisteita. PAS-palvelut kirjaavat hyödyntävän organisaation ohjelmiston toimenpiteet niin, että jokaisesta HTTP-kyselyviestistä tallentuu vähintään seuraavat asiat automaattisesti:

- Kuka teki? (käyttäjätunnus, jolla viesti lähetettiin)
- Mitä teki? (kyselyn sisältö)
- Milloin teki? (kyselyn aikaleima)
- Missä teki? (IP-osoite, josta viesti lähetettiin)
- Mitä vastattiin? (HTTP-statuskoodi ja vastauksen koko tavuina)

Autentikointi edellyttää käyttäjätunnusta ja salasanaa. Näiden käyttöä ei tässä dokumentissa kuvata, koska se riippuu hyödyntävän organisaation ohjelmiston teknisestä toteutuksesta.

### 2.2 Viestien rakenne

#### 2.2.1 Kyselyviestit

Kyselyviestit ovat muotoa:

```
{GET, HEAD, PATCH, POST, DELETE} <base>/<contract>/<term>...
```

Näistä `<term>` kuvaa olemassa olevaa resurssia tai työkalua, ja se on jokin seuraavista:

- `transfers` – Siirtojen resurssi.
- `search` – Hakutyökalu.
- `preserved` – Säilytyspakettien resurssi.
- `disseminated` – Jakelupakettien resurssi.
- `statistics` – Tilastot.
- `history` – Historiatietojen resurssi.
- `public-keys` – Julkisten avainten resurssi.

Kyselyviestien muoto on tässä määrittelyssä esitetty ilman vapaaehtoisia HTTP-otsakekenttiä, joita voi lisätä mihin tahansa viestiin RFC 7231 -määrittelyn [RFC\_7231] mukaisesti. Esimerkiksi viestiin:

```
GET https://pas.csc.fi/api/2.0/csc
```

voidaan sisällyttää kenttiä seuraavasti:

```
GET https://pas.csc.fi/api/2.0/csc HTTP/1.1
Date: Mon, 13 Nov 2017 08:12:31 GMT
```

```
Host: pas.csc.fi
...
```

POST-kutsuissa lähetävä datan enkoodauksessa käytetään `application/offset+octet-stream` -muotoa.

### 2.2.2 Palautusviestit

Palautusviestien rungossa käytetään JSON-muotoa [RFC\_7159] käyttäen JSend-rakennetta [JSEND]. Onnistuneessa tilanteessa palautusviestiksi annetaan seuraava rakenne:

```
{
  "status": "success",
  "data": { ... }
}
```

Palautusviesteissä `status` kertoo kyselyn onnistumisesta tai epäonnistumisesta, ja `data` tulosjoukon. Epäonnistuneessa tilanteessa vastausviestiksi annetaan seuraava rakenne:

```
{
  "status": "fail",
  "data": { "message": "..."}
}
```

Virheilanteessa `message` kertoo virheilmoituksen. Jos kuitenkin kyselyn parametri on ollut virheellinen, palautetaan virheilmoitus parametrin nimellä:

```
{
  "status": "fail",
  "data": { "<parameter>": "..."}
}
```

Esimerkki: Jos kyselyssä on ollut parametri `limit`, ja sille on annettu vääränlainen arvo, voisi ilmoitus olla seuraava:

```
{
  "status": "fail",
  "data": { "limit": "Value can only be an integer in range 1-1000"}
}
```

Palautettava säilytyksestä haettu aineisto on ZIP- tai TAR-jakelupaketti, tai XML-muotoinen metadatatodokumentti.

## 2.3 Rajapinnan toiminnot

Resurssi	Metodit	Kuvaus
<code>/transfers</code>	POST	Aineiston siirron alustaminen
<code>/transfers/&lt;siirron tunniste&gt;</code>	HEAD, PATCH	Aineiston siirto PAS-palveluun
<code>/transfers</code>	GET	Siirtojen listaus
<code>/transfers/&lt;siirron tunniste&gt;/status</code>	GET	Siirron tila PAS-palvelun vastaanotossa
<code>/transfers/&lt;siirron tunniste&gt;/report</code>	GET	Siirtopaketin tarkastusraportin nouto
<code>/search</code>	GET	Hakutoiminto
<code>/preserved/&lt;aip-id&gt;</code>	GET	Säilytyspaketin tiedot PAS-palvelussa
<code>/preserved/disseminate</code>	POST	Säilytettävän aineiston muodostaminen jakelupaketiksi
<code>/disseminated</code>	GET	Jakelupakettien listaus
<code>/disseminated/&lt;dip-id&gt;</code>	GET	Jakelupaketin tilan PAS-palvelussa
<code>/disseminated/&lt;dip-id&gt;/download</code>	GET	Jakelupaketin nouto
<code>/history/transfers</code>	GET	Siirtojen historiatietojen listaus



Resurssi	Metodit	Kuvaus
/history/preserved	GET	Säilytettävän aineiston historiatietojen listaus
/history/disseminated	GET	Aineiston jakelun historiatietojen listaus
/statistics/overview	GET	Tilastotiedot
/public-key/dip	GET	Jakelutoiminnon julkisen avaimen nouto

## 2.4 Estetyt tasot

Seuraavien viestitasojen käyttö on tarkoituksella estetty.

### Viestin muoto:

```
GET <base>  
GET <base>/<contract>  
GET <base>/<contract>/preserved  
GET <base>/<contract>/disseminated  
GET <base>/<contract>/ingest  
GET <base>/<contract>/ingest/report  
GET <base>/<contract>/statistics  
GET <base>/public-key
```

## 3 AINEISTON SIIRTO

Aineistojen siirtämisessä PAS-palveluun hyödyntävän organisaation ohjelmisto lähettää siirtopaketteja PAS-palveluun, seuraa vastaanoton tilaa, ja saa siirtopaketeista myöhemmin tarkastusraportteja käsiteltäväkseen. Siirtopaketti voidaan joko hyväksyä säilytykseen tai hylätä PAS-vastaanotossa.

Aineistojen siirtämiseen PAS-palveluihin käytetään TUS-protokollaa [TUS]. TUS-protokolla on luotu isojen tiedostojen siirtoon ja se on rakennettu HTTP protokollan päälle. Protokolla toimii pilkkomalla ja lähettämällä tiedoston pienissä osissa, jolloin tiedoston lähettäminen voidaan keskeyttää ja jatkaa myöhemässä vaiheessa. Aineistojen lähettäminen TUS-protokollan avulla vaatii tyypillisesti sovelluksen, asiakasohjelman, joka asennetaan lähettävään päähän.

Aineiston siirrossa luodaan jokaiselle siirtotapahtumalle yksilöllinen siirron tunniste (transfer-id), jonka avulla siirtoa voi seurata. Siirron tunniste ei ole sama kuin siirtopaketin tunniste. Jos samaa siirtopakettia lähettää PAS-palvelun monta kertaa, jokainen siirto saa oman tunnisteensa siirron alustuksessa. Tämän avulla käyttäjä voi seurata ja ladata itselleen jokaisen yksittäisen siirron siirtopaketin tarkastusraportit.

Siirron voi keskeyttää milloin tahansa ja jatkaa myöhemmin. TUS-protokollan asiakasohjelmistojen avulla voidaan tietää, kuinka suuri osa tiedostosta on jo siirretty ja mistä kohdasta tiedoston siirtämistä voi jatkaa. Jos siirron haluaa keskeyttää ja perua kokonaan siirtämisen aikana, voi yksinkertaisesti vain keskeyttää siirron. PAS-palvelu siivoaa keskeneräiset hylätyt siirrot tietyn ajan kuluttua eikä vastaanoton työvuotoa käsiteltäväkseen keskeneräistä siirtoa.

Aineiston siirto on aktiivinen 20 päivän ajan, jonka aikana siirtopaketin tarkastusraportit ovat haettavissa PAS-palvelusta. 20 päivän jälkeen tarkastusraportit poistetaan ja siirron tunniste, transfer-id, ei ole enää haettavissa. Vanhoja siirtotapahtumia voi listata historiakomentojen avulla.

### 3.1.1 Lähetyksen alustaminen

Siirtopakettien lähetykset on alustettava lähettämällä siirtopaketin metatiedot palvelulle (TUS-protokollan "Upload Creation"). Metatiedoissa kerrotaan mm. lähettävän tiedoston koko ja luotavan tiedoston nimi. Palautusviestissä annetaan siirron yksilöivä tunniste (transfer-id) ja osoite, johon siirtopakettia voi alkaa lähettämään. Tiedoston kokoa käytetään, kun tiedostoa lähetetään osissa ja sen avulla ohjelmisto tietää milloin tiedosto on siirretty kokonaisuudessaan PAS-palvelun vastaanottoon.

#### Viestin muoto:

```
POST <base>/transfers
```

#### Viestin määreet:

Viestissä annetaan ainoastaan tietoa pyynnön otsikossa. Pynnön otsikossa on annettava seuraavat tiedot avain-arvo -pareina:

- Tus-Resumable – TUS-protokollan versio.
- Upload-Length – Siirtopakettitiedoston koko tavuina.

### 3.1.2 Tiedoston lähettäminen

Kun lähetyksen alustettu onnistuneesti, osoite, johon siirtopakettia voi alkaa lähettämään, on tiedossa. Paketin lähettäminen tapahtuu PATCH metodilla. TUS-protokolla pilkkoo tiedoston osiin ja lähettää osat PAS-palveluun. Siirtopaketin lähettäminen tapahtuu osissa, lähettämällä useita PATCH-pyyntöjä, kunnes koko tiedosto on lähetetty. Tyypillisesti TUS-asiakasohjelma hoitaa tiedoston pilkkomisen osiin ja osien lähettämisen.

#### Viestin muoto:

```
PATCH <base>/transfers/<transfer-id>
```

#### Viestin määreet:

- `transfer-id` – Siirron yksilöivä tunniste.

Viestin rungossa annetaan siirtopaketti osina `application/octet-stream` mediatyyppinä.

Lisäksi pyynnön otsikoissa on annettava seuraavat tiedot:

- `Tus-Resumable` – TUS-protokollan versio
- `Upload-Offset` – Viestin rungossa lähetettävän osan etäisyys tavuina paketin alusta. Tämä kertoo mihin kohtaan tiedostoa, lähetyksen osa kuuluu. Jos lähetyksen keskeytetään, tämän arvon avulla voidaan jatkaa lähetystä siitä mistä jäättiin.

Kun `Upload-Offset` vastaa lähetysten alustamisessa annettua siirtopaketin tiedoston kokoa, lähetyksen on valmis ja siirtopaketti on kokonaisuudessaan siirretty onnistuneesti PAS-palveluun.

### 3.1.3 Lähetysten tilan kysely

Kun lähetyksen on alustettu onnistuneesti, osoite, josta lähetysten tilaa voi kysellä, on tiedossa. Tilan kysely tapahtuu `HEAD` metodilla. Vastauksessa ilmenee lähetysten vastaanotettu koko palvelimella, jolloin loppukäyttäjä tietää kuinka suuren osan lähetyksestä on jo lähetetty, sekä alustuksessa annetut metatiedot ja lähetysten lopullisen koon. TUS-asiakasohjelmistot käyttävät tätä toimintoa hyväkseen, kun siirtoa on aikaisemmin keskeytetty ja nyt halutaan jatkaa tiedoston siirtämistä.

**Viestin muoto:**

```
HEAD <base>/transfers/<transfer-id>
```

**Viestin määreet:**

- `transfer-id` – Siirron yksilöivä tunniste.

Viestissä runko on tyhjä. Pynnön otsikoissa on annettava seuraavat tiedot:

- `Tus-Resumable` – TUS-protokollan versio

### 3.1.4 Siirtojen luettelointi

Tämän toiminnon avulla käyttäjä voi luetella kaikki aktiiviset siirrot. Vastauksessa palautetaan tulosjoukko siirroista ja tulosalkioissa osoite, jonka avulla siirron tarkemmat tiedot ovat saatavilla. Lisäksi siirtojen aikaleimat ja tilat näytetään osana tulosalkiota. Kun siirto on yli 20 päivää vanha, sen tiedot poistetaan eikä se ole enää luettavissa tämän toiminnon avulla.

**Viestin muoto:**

```
GET <base>/transfers
```

**Viestin määreet:**

- `status` – Tämän avulla tulosjoukkoa voi suodattaa siirron tilan mukaan.

### 3.1.5 Vastaanoton tilan esittäminen

Kun siirto on alustettu ja käyttäjä on saanut vastauksessa siirron tunnisteeseen, voi siirron tilaa kysellä tämän toiminnon avulla. Vastauksessa esitetään siirron tila PAS-palvelun vastaanotossa. Tilat ovat:

- hyväksytyt – siirtopaketti on hyväksytyt pitkäaikaissäilytykseen ja PAS-palvelu on ottanut säilytysvastuun aineistosta
- hylätty – siirtopaketissa on virheitä ja se on hylätty PAS-palvelun vastaanotossa
- käsittelyssä – siirtopaketti on siirretty kokonaisuudessaan ja PAS-palvelun vastaanoton tarkistukset ovat käynnissä
- siirto kesken – siirtopaketin siirtoa on aloitettu, mutta siirtäminen on vielä kesken eikä PAS-palvelun vastaanotto ole vielä ottanut siirtopakettia käsiteltäväkseen

Mikäli siirto on hyväksytty tai hylätty, vastauksessa ilmoitetaan myös osoite, josta siirtopaketin tarkastusraportit ovat ladattavissa.

**Viestin muoto:**

```
GET <base>/transfers/<transfer-id>/status
```

**Viestin määreet:**

- `transfer-id` – Siirron yksilöivä tunniste.

**3.1.6 Siirtopaketin tarkastusraportin nouto**

Kun siirto on käsitelty PAS-palvelun vastaanotossa, on siirtopaketin tarkastusraportti noudettavissa tämän toiminnon avulla.

**Viestin muoto:**

```
GET <base>/transfers/<transfer-id>/report
```

**Viestin määreet:**

- `transfer-id` – Siirron yksilöivä tunniste.

## 4 AINEISTON HAKU

Säilytyksessä olevat aineistot ovat löydettävissä aineiston mukana PAS-palveluun vastaanotettujen metatietojen avulla. Hakutoiminnolla voidaan muun muassa selata säilytyksessä olevaa aineistoa ja hakea tunnisteita jakelupyynnöitä varten.

### 4.1 Hakuviestit

Käsiteltävää aineistoa voi rajata halutuksi asettamalla ehtoja koskien aineistojen metatietoja. PAS-järjestelmässä tietty aineisto tallennetaan yhteen tai useampaan säilytyspakettiin. Hakutoiminnolla voi etsiä säilytys- ja jakelupaketteja ja saada tuloksena luettelon niistä tietopaketeista, joissa annettu hakuehto toteutuu. Jokainen säilytyspaketti sisältää paketin sisällön kuvaavan METS-dokumentin, jonka tiedot ovat haettavissa PAS-palvelujen hakutoiminnolla. Säilytyksessä olevan aineiston päivittäminen (esim. migraatio tai *Aineistojen ja niiden metatietojen paketointi pitkäaikaissäilytykseen* -määrittelyn [FI\_META] mukainen päivityspaketti) ei muuta jo olemassa olevia säilytyspaketteja, vaan päivitetuille versioille luodaan uudet säilytyspaketit. Näin hakutoiminnolla ja sen palauttamilla tunnisteilla on mahdollista erottaa myös aineistojen eri versiot toisistaan.

#### Viestin muoto:

```
GET <base>/<contract>/search?<parameters>
```

#### Viestin määreet:

- `<parameters>` – Kyselyssä voi käyttää alla olevia parametreja `&`-merkeillä erotettuna hakutulosten rajaamiseksi.

```
q=<query>
```

Hakuehto, jota käytetään tulosjoukon haussa. Hakuehtoa esitellään erikseen luvussa 4.2. Oletuksena haetaan kaikesta aineistosta. On huomattava, että muodostettu hakuehto tulee URL-koodata REST-rajapintaa varten.

```
limit=<limit>
```

Parametrilla voi määrittellä palautettavien tulosten enimmäismäärän. Oletuksena käytetään arvoa 20. Kokonaisluvut väliltä 1-1000 ovat sallittuja.

```
page=<page>
```

Mikäli hakutuloksia on enemmän kuin kerralla palautetaan, `<page>`-parametria käyttäen voidaan näyttää myöhempiä hakutuloksia. Esimerkiksi 50 hakutuloksesta `page=2` palauttaa hakutulokset 21–40 käytettäessä `limit`-oletusarvoa 20 tulosta/sivu. Oletuksena käytetään arvoa 1, joka vastaa ensimmäistä sivua.

### 4.2 Hakutoiminto

Säilytyspakettien ja olemassa olevien jakelupakettien METS-dokumentit on kokonaan haettavissa METS-tietokannasta, ja niiden tietoja voidaan hakea tässä luvussa esitetyllä tavalla. Hakuehdossa käytetään Apache Lucene –syntaksia [LUCENE]. Hakuehdossa käytetään aina avain-arvo-pareja muodossa `'avain:arvo'`, jossa avain vastaa METS-dokumentin tiettyä elementtiä tai attribuuttia ja arvo tämän avaimen sisältämää arvoa.

Säilytys- ja jakelupaketit ovat haettavissa, kun niiden tiedot ovat indeksoituneet hakutietokantaan. Jos säilytyspaketin säilytysvastuu on vasta äskettäin siirtynyt tai jakelupaketti on vasta äskettäin valmistunut, se ei välttämättä ole vielä indeksoitunut tietokantaan eikä pakettia tällöin hakutoiminnolla löydy. Samoin, kun jakelupaketti poistetaan, sen tiedot siivotaan hakutietokannasta viiveellä. Tällä viiveellä ei kuitenkaan ole vaikutusta muihin rajapinnan toimintoihin ellei erikseen mainita.

#### 4.2.1 Hakutoiminnon ominaisuudet ja syntaksi

Toiminnot ovat hakutoiminnossa käytettävissä Apache Lucene –syntaksilla [LUCENE] luqum-sovelluskirjaston [LUQUM] mukaisesti ja se sisältää mm. seuraavia toiminnallisuuksia<sup>1</sup>:

- loogiset operaattorit AND, OR ja NOT, joita voidaan myös sulkerkein ryhmitellä
- sumea haku (fuzzy search)
- läheisyshaku (proximity search)
- jokerimerkit ? ja \* (wildcard)
- haku tietyltä alueelta (range search)
- termin merkityksellisyyden korostaminen (boosting a term)

#### 4.2.2 Hakulauseet

Hakuehdossa avaimen nimi vastaa METS-dokumentin tiettyä elementtiä tai attribuuttia. Arvo taas jotakin, mitä tämän elementin tai attribuutin tulee sisältää. Avaimena voi käyttää myös XML-polkua, jossa elementtien ja attribuuttien nimet erotellaan alaviivalla ilman nimiavaruuksia. Koko polkua ei kuitenkaan ole pakko antaa vaan pelkkä loppuosa riittää. Avaimet ovat merkkikokoriippuvaisia, mutta arvoja käsitellään merkkikokoriippumattomasti. Esimerkkejä<sup>2</sup>:

- `q='mets_OBJID:id123'` – Hae tiedot paketeista, joiden METS-dokumentissa on sellainen `mets-elementti`, jonka `OBJID-attribuutti` sisältää arvon `id123`. Esimerkissä avaimena toimii `mets_OBJID`.
- `q='subject:xxx'` – Hae tiedot paketeista, joiden METS-dokumentissa mikä tahansa `subject-elementti` tai `-attribuutti` sisältää arvon `xyz`. Tuetuista kuvailevan metatiedon muodoista tämä huomioi Dublin Coren lisäksi myös esimerkiksi EAD- ja MODS-muotojen `subject-elementit`.
  - Koska täyttää polkua ei anneta, saa hakuavain olla METS-dokumentissa mikä tahansa.
  - Esim. `q='mets_dmdSec_mdWrap_xmlData_subject:xyz'` edellyttää arvon `xyz` sijaintia seuraavassa avainpolussa: `<mets>/<dmdSec>/<mdWrap>/<xmlData>/<subject>`. Nimiavaruus saa METS-dokumentin XML-rakenteessa vaihtua kesken polun.
- `q='objectIdentifierValue:xyz'` – Hae tiedot paketeista, jotka sisältävät digitaalisen objektin, jonka tunnisteen arvo on `xxx`.
- `q='formatName:image/tiff'` – Hae tiedot paketeista, jotka sisältävät TIFF-kuvia.
- `q='MDTYPE:DC'` – Hae tiedot paketeista, jotka sisältävät Dublin Core -metadattaa.
- `q='author:ki* AND title:seitse*'` – Hae tiedot paketeista, jonka metatiedoissa `author-elementti` tai `-attribuutti` alkaa `ki`-jonolla, ja `title-elementti` tai `-attribuutti` `seitse`-jonolla.

<sup>1</sup> Katso ohjeet ja esimerkit hakuehdon syntaksin käyttämiseksi osoitteesta:

[https://lucene.apache.org/core/3\\_6\\_0/queryparsersyntax.html](https://lucene.apache.org/core/3_6_0/queryparsersyntax.html)

<sup>2</sup> Lisää esimerkkejä julkaistaan osoitteesta: <https://www.digitalpreservation.fi/specifications/examples>

## 5 SÄILYTETTÄVÄN AINEISTON HALLINTA

Rajapinnan avulla voidaan hallita kokonaisia säilytyspaketteja. Useita säilytyspaketteja voidaan hallita niin, että hyödyntävän organisaation ohjelmisto käy ne automaattisesti halutulla tavalla läpi. Tällainen läpikäynti on tarpeen, jos esimerkiksi halutaan hallita kaikkia niitä paketteja, jotka saatiin luettelona luvun 3 mukaisella hakuehdolla.

Deleted: 3.3

### 5.1 Säilytyspaketin tiedot

Tällä toiminnolla saa tietoja säilytyspaketista. Palautusviestissä näkyy säilytyspaketin yleisiä tietoja kuten päivämääriä ja METS-dokumentin tunniste, mutta myös säilytyspaketin sisältämien tiedostojen tunnisteet ja niiden rakenne. Toiminnon palautustietojen avulla voidaan muodostaa säilytyspaketin sisällöstä aineiston jakelupyynnö.

#### Viestin muoto:

```
GET <base>/<contract>/preserved/<aip-id>
```

#### Viestin määreet:

- `<aip-id>` – Säilytyspaketin tunniste, jonka perusteella kysely rajataan koskemaan säilytyksessä olevaa aineistoa. Tunnisteita voi hakea luvun 3 mukaisella toiminnolla.

Deleted: 3.3

### 5.2 Aineiston muodostus jakelupaketiksi

Tällä toiminnolla voi muodostaa jakelupaketin halutusta aineistosta, joka on pitkäaikaissäilytyksessä. Toiminto muodostaa hyödyntävän organisaation ohjelmiston saataville aina uuden jakelupaketin omalla tunnisteellaan, eikä se poista aiemmin luotuja jakelupaketteja. Jakelupaketeista on kerrottu luvussa 6.

#### Viestin muoto:

```
POST <base>/<contract>/preserved/disseminate
```

#### Viestin määreet:

Viestiin liitetään JSON-tiedosto, joka kertoo mistä aineistosta jakelupaketti muodostetaan. Jakelupakettiin halutun aineiston yksilöintiin käytetään sopimustunnuksen lisäksi säilytyspakettien, rakennekarttojen tai digitaalisten objektien tunnisteita. Viestiin liitettävän JSON-tiedoston täytyy sisältää ainakin jokin seuraavista tiedoista, jotka ovat toistettavissa:

- `<aip_id>` – Säilytyspaketin tunniste, josta jakelupaketti muodostetaan. Jos annetaan säilytyspaketin tunniste kokonaisuudessaan, muodostetaan jakelupaketti kokonaisesta säilytyspaketista.
- `<aip-id:file-id>` – Säilytyspaketissa olevan tiedoston tunniste. Koska tiedoston tunniste ei ole yksilöllinen tietopaketin ulkopuolella ja on siten annettava yhdessä säilytyspaketin tunnisteiden kanssa.
- `<aip-id:div-id>` – Rakennekartan tunniste voi palauttaa yhden tai useamman digitaalisen objektin. Rakennekartan tunniste ei ole yksilöllinen tietopaketin ulkopuolella ja on siten annettava yhdessä säilytyspaketin tunnisteiden kanssa.

JSON-tiedostossa voidaan lisäksi antaa seuraavia tietoja:

- `<dip-name>` – Jakelupakettitiedostolle annettava nimi
- `<metadata>` – Jakelupaketti muodostetaan pyydetyin aineiston pelkistä metatiedoista ilman digitaalisia objekteja.
- `<catalog>` – Jakelupaketti muodostetaan skeemakatalogin version `<catalog>` mukaisesti. Parametrina annetaan versionumerosta vain kaksi ensimmäistä numeroa. Esim. `catalog=1.6` palauttaa tulokset skeemakatalogin 1.6.X mukaisesti, missä X on uusin olemassa oleva 1.6-

versiosarjan numero. Jakelupakettia ei voi muodostaa säilytyspakettia vanhempaan katalogiversioon. Oletuksena käytetään uusinta versiota.

- `<format>` - Jakelupaketti käännetään tiedostomuotoon `<format>`. Sallitut arvot ovat `zip` ja `tar`. Oletuksena käytetään ZIP-pakkausta. TAR-muoto palautetaan ilman pakkausta.



## 6 AINEISTON JAKELU

Kun PAS-järjestelmästä halutaan noutaa aineistoa, pitää halutusta säilytyksessä olevasta aineistosta ensin muodostaa jakelupaketti (ks. luku 5.2), jonka voi noutaa. PAS-palvelu muodostaa jakelupaketin *Aineistojen ja niiden metatietojen paketointi pitkäaikaissäilytykseen* -määrittelyn [FI\_META] mukaiseksi. Paketille annetaan uusi tunniste, ja sille muodostetaan METS-dokumentti. Lopuksi vielä paketin eheys varmistetaan, ja sen jälkeen se asetetaan saataville. Aineiston nouto ei poista aineistoa säilytyksestä. Jakelupaketti on noudettavissa PAS-palvelusta kymmenen päivän ajan.

Deleted: 3.4.2

### 6.1 Muodostettavien ja noudettavien jakelupakettien listaus

Tällä toiminnolla listataan sellaisten jakelupakettien tunnisteet, joiden muodostus on kesken tai jotka ovat noudettavissa. Samalla toiminnolla saa myös luettelon niistä komennoista, jotka voidaan kohdistaa jakelupaketille.

**Viestin muoto:**

```
GET <base>/<contract>/disseminated
```

### 6.2 Jakelupaketin seuranta ja niihin kohdistuvat komennot

Tällä toiminnolla voi seurata jakelupaketin muodostumista. Samalla toiminnolla saa myös luettelon niistä komennoista, jotka voidaan kohdistaa jakelupaketille.

**Viestin muoto:**

```
GET <base>/<contract>/disseminated/<dip-id>
```

**Viestin määreet:**

- `<dip-id>` – Jakelupaketin tunniste, joka löytyy jakelupaketin muodostusviestissä luvun 5.2 kyselyssä. Muodostettuja jakelupaketteja voi selata luvun 4 mukaisesti.

### 6.3 Muodostettujen jakelupakettien noutaminen

Tällä toiminnolla voidaan noutaa muodostettuja jakelupaketteja tai niiden metatietoja. Muodostetut jakelupaketit säilyvät PAS-palvelussa kymmenen (10) päivää.

**Viestin muoto:**

```
GET <base>/<contract>/disseminated/<dip-id>/download
```

## 7 KÄSITTELYHISTORIA

Jokaisesta siirrosta, säilytyspaketista ja muodostetusta jakelupaketista tallennetaan tieto käsittelyhistoriaan. Tämä resurssi auttaa PAS-palveluiden käyttäjiä löytämään aineistonsa ja toimii lokitietona hyväksytyistä ja hylätyistä siirtopaketeista ja muodostetuista jakelupaketeista, joiden elinkaari PAS-palvelussa on määräaikainen. Käsittelyhistoriaan tallennetaan tyypillisesti tietopakettien tunnisteita ja aikaleimoja.

Tietopaketit määrittelevät säilytettävän tai jaeltavan kokonaisuuden. Yksi siirto- tai säilytyspaketti sisältää yleensä useita tiedostoja, jotka muodostavat yhden säilytettävän kokonaisuuden. Säilytettävä kokonaisuus määritellään säilytyksen aloittamista varten. Jakelupaketti eli säilytyksestä noudettava aineisto voi sisältää tiedostoja useammasta kuin yhdestä säilytettävästä kokonaisuudesta eli säilytyspaketista.

### 7.1 Siirtohistoria

Tämä toiminto listaa vastaanotossa tarkastetut siirtopaketit ja niiden siirtopäivämäärän, tietopaketin nimen ja tarkastuksen tuloksen.

**Viestin muoto:**

```
GET <base>/<contract>/history/transfers
```

**Viestin määreet:**

- Suodatin – Siirtohistoriaa voi selata vuoden tai kuukauden tarkkuudella.

### 7.2 Säilytyshistoria

Tämä toiminto listaa säilytykseen hyväksytyt säilytyspaketit ja niiden säilytyksen aloituksen ajankohdan, tietopaketin nimen, joka on sama kuin siirtopaketilla, säilytyspaketin tunnisteen, METS-dokumentin tunnisteen ja säilytyspaketin koon.

**Viestin muoto:**

```
GET <base>/<contract>/history/preserved
```

**Viestin määreet:**

- Suodatin – Säilytyshistoriaa voi selata vuoden tai kuukauden tarkkuudella.

### 7.3 Jakeluhistoria

Tämä toiminto listaa säilytyksessä olevasta aineistosta muodostetut jakelupaketit ja niiden valmistuspäivämäärän, tietopaketin nimen, säilytyspaketin tunnisteen ja jakelupaketin muodostajan.

**Viestin muoto:**

```
GET <base>/<contract>/history/transfers
```

**Viestin määreet:**

- Suodatin – Jakeluhistoriaa voi selata vuoden tai kuukauden tarkkuudella.

## 8 TILASTOT

PAS-palvelun käytöstä on saatavilla sopimuskohtaista tilastotietoa. Tätä osiota tullaan tulevaisuudessa kehittämään eteenpäin. Myöhemmin lisätään mahdollisuus hakea tilastoja tietyllä ajanjaksolla. Tilastojen muodostamisessa voi esiintyä viivettä.

### 8.1 Yhteenveto tilastoista

Tämä toiminto antaa yhteenvedon sopimukselle kohdistuvista tilastoista.

**Viestin muoto:**

```
GET <base>/<contract>/statistics/overview
```

## 9 JULKISET AVAIMET

Julkinen avain on käyttäjän, aineiston tai PAS-palvelun tunnistamisessa ja todentamisessa käytetyn avainparin julkinen osa. Julkisella avaimella voidaan todentaa siirretyn aineiston lähettäjä ja eheys digitaalisen allekirjoituksen avulla.

### 9.1 PAS-palvelun jakeluavaimen noutaminen

Toiminnolla voi noutaa jakelupaketin digitaalisessa allekirjoituksessa käytetyn avainparin julkisen osan. Avainta voi käyttää sen todentamiseen, että jakelupaketti on tullut PAS-palvelusta ja että se on eheä.

**Viestin muoto:**

```
GET <base>/public-key/dip
```

## LIITE A. ESIMERKKEJÄ PROSESSIEN AUTOMATISOINNISTA

Automaattisina esimerkkiprosesseina luvuissa Q ja Q esitetään aineistojen siirtäminen PAS-palveluun ja aineistojen noutaminen PAS-palvelusta. Nämä esimerkit eivät kuitenkaan rajoita rajapinnan hyödyntämistä toisella tavalla: saman lopputuloksen tuottavat automaatioprosessit voidaan toteuttaa monella muullakin tavalla. Esitetyt esimerkkiprosessit eivät ole yksityiskohtaisia, vaan esitystapa on tässä pyritty pitämään kompaktina ja johdattelevana.

Deleted: A.1

Deleted: A.2

### Automaatioprosessi aineistojen siirtämiseksi PAS-palveluun

Aineistojen siirtämisessä PAS-palveluun hyödyntävän organisaation ohjelmisto lähettää siirtopaketteja PAS-palveluun ja saa niistä myöhemmin tarkastusraportteja käsiteltäväkseen. Siirtopaketti voidaan joko hyväksyä säilytykseen tai hylätä PAS-vastaanotossa. Prosessina voisi olla esimerkiksi seuraava komentosarja, jonka hyödyntävän organisaation ohjelmisto tekee automaattisesti päivittäin:

Hyödyntävän organisaation ohjelmisto...

1. valitsee päivän aikana paketoitavan aineistokokonaisuuden.
2. paketoii valitun aineiston paketoitimäärityksen [FI\_META, FI\_FORMATS] mukaisiksi siirtopaketeiksi.
3. avaa yhteyden kirjautumalla PAS-palveluun.
4. toistaa seuraavat komennot jokaiselle edeltävänä päivänä saapuneelle tarkastusraportille koskien hyväksytyjä paketteja:
  - a. noutaa tarkastusraportin PAS-palvelusta.
  - b. käsittelee tarkastusraportin hyödyntävässä organisaatiossa määritellyllä tavalla.
  - c. poistaa tarkastusraportin PAS-palvelusta.
  - d. tarvittaessa poistaa tarkastusraporttia koskevan siirtopaketin ja sen sisältämän aineiston hyödyntävän organisaation järjestelmistä (aineisto on tämän jälkeen vain pitkäaikaissäilytyksessä).
5. toistaa seuraavat komennot jokaiselle edeltävänä päivänä saapuneelle tarkastusraportille koskien hylättyjä paketteja:
  - a. noutaa tarkastusraportin ja tarvittaessa hylätyn siirtopaketin PAS-palvelusta.
  - b. käsittelee tarkastusraportin hyödyntävässä organisaatiossa määritellyllä tavalla.
  - c. voi tässä vaiheessa poistaa tarkastusraportin ja hylätyn siirtopaketin PAS-palvelusta, mutta jos hylkäyksen syyn selvittämiseen tarvitaan PAS-palvelun apua, on hyödyllisempää jättää tarkastusraportti ja hylätty paketti PAS-palveluun.
6. toistaa seuraavat komennot jokaiselle prosessin vaiheessa 2 luodulle siirtopaketille:
  - a. siirtää siirtopaketin PAS-palveluun.
7. sulkee yhteyden PAS-palveluun.

Edellä olevassa esimerkkiprosessissa noudetaan edeltävänä päivänä syntyneet tarkastusraportit ennen uusien siirtojen aloittamista, jonka jälkeen siirretään päivän aineisto PAS-palveluun. Näin prosessi voidaan ajastaa automaattisesti suoritettavaksi kerran vuorokaudessa. Prosessi voidaan kuitenkin automatisoida muullakin tavalla. Mikään ei esimerkiksi velvoita odottamaan valmistuneita tarkastusraportteja seuraavaan päivään, vaan ne voidaan yhtä hyvin noutaa heti valmistuttuaan.

### Automaatioprosessi aineistojen noutamiseksi PAS-palvelusta

Hyödyntävän organisaation ohjelmisto voi muodostaa PAS-palvelussa olevasta aineistosta jakelupaketteja ja noutaa ne. Noudettava aineistoa voi rajata valintakriteeristöillä, eli aineiston metatietoja koskevilla ehdoilla. Valintakriteeristö rajaa aineistoa ja muodostaa näin halutun aineistokokonaisuuden, joka koostuu yhdestä tai useammasta säilytyspaketistä. Aineistojen noudon automaatiologiikan toteuttamisen helpottamiseksi PAS-palvelu tarjoaa vastauksissaan suoraan ne komennot, joita aineiston noutoprosessin seuraavassa vaiheessa tarvitaan. Näin ollen prosessissa eteenpäin pääsemiseksi tavallisesti riittää, että hyödyntävän organisaation ohjelmisto poimii saaduista vastauksista nämä komennot ja välittää ne muuttamattomina takaisin PAS-palvelulle kyselyinä. Prosessina voisi olla esimerkiksi seuraava komentosarja, jonka hyödyntävän organisaation ohjelmisto tekee automaattisesti:

Hyödyntävän organisaation ohjelmisto...

1. avaa yhteyden kirjautumalla PAS-palveluun.

2. luo (tai lukee paikallisesti) valintakriteeristön, joka rajaa noudettavan aineiston halutuksi. Valintakriteeristö on joukko ehtoja, jotka koskevat halutun aineiston metatietoja. Ks. lisätietoa kriteeristön muodostamisesta luvusta [4.2](#).
3. lähettää valintakriteeristön PAS-palveluun. PAS-palvelu toimittaa vastauksena luettelon kaikista valintakriteeristön täyttävien säilytyspakettien tunnisteista (ja jonkin verran muita metatietoja). Valintakriteeristö täyttyy säilytyspaketille, jos sen sisältämän aineiston metatiedot vastaavat kriteeristössä esitettyjä vaatimuksia. Tunnisteiden mukana tulee myös näiden säilytyspakettien osoitteet, jotka toimivat rajapinnassa suoraan komentoina säilytyspakettien hallitsemiseksi. Ks. lisätietoa luvusta [3](#).
4. toistaa seuraavat komennot jokaiselle PAS-palvelusta vaiheessa 3 saadulle tunnisteelle:
  - a. komentaa PAS-palvelua muodostamaan tunnisteiden omaavasta säilytyspaketista jakelupaketin hyödyntäen tunnisteiden yhteydessä saatua osoitetta. PAS-palvelu lähettää kuittauksena tulevan jakelupaketin osoitteen. Tämä osoite toimii rajapinnassa suoraan komentona jakelupaketin hallitsemiseksi. Ks. lisätietoa luvusta [5](#).
  - b. seuraa jakelupaketin osoitteen avulla jakelupaketin muodostumista ja odottaa, kunnes jakelupaketti on valmis. Kun jakelupaketti on valmis, PAS-palvelu palauttaa vastauksessa ohjelmistolle osoitteen, joka toimii rajapinnassa suoraan komentona jakelupaketin noutamiseksi. Ks. lisätietoa luvusta [6.1](#).
  - c. noutaa jakelupaketin saadun osoitteen avulla. Ks. lisätietoa luvusta [6.3](#).
  - d. varmistaa, että saatu jakelupaketti on noudossa säilynyt eheänä. Ohjelmisto toistaa kohdat c ja d, jos ei ole. Näitä vaiheita voidaan toistaa, kunnes jakelupaketti on todettu saapuneen eheänä.
  - e. poistaa jakelupaketin PAS-palvelusta. Ks. lisätietoa luvusta [Error! Reference source not found.](#)
5. sulkee yhteyden PAS-palveluun.

Deleted: 3.3.1

Deleted: 3.3

Deleted: 3.4

Deleted: 3.5.1

Deleted: 3.5.2

Deleted: 3.5.3

Prosessi voidaan automatisoida muullakin tavalla. Esimerkiksi kokonaisia valintakriteeristöjä voisi olla yhdessä prosessiajossa useita erilaisia, jolloin vaiheet 2 – 4 voi automaattisesti toistaa jokaiselle kriteeristölle erikseen. Tällainen tilanne voi olla esimerkiksi silloin, jos käytettävissä on luettelo tunnisteista, ja halutaan noutaa kutakin tunnistetta vastaavat aineistot. Prosessia on myös mahdollista tehostaa esimerkiksi niin, että seuraavan jakelupaketin muodostus komennetaan aloitettavaksi jo ennen käsittelyssä olevan jakelupaketin noutoa. Tällöin seuraava jakelupaketti on PAS-palvelussa muodostumassa samaan aikaan, kun käsiteltävää jakelupakettia noudetaan ja sen eheyttä varmistetaan. Lisäksi vaihe 4.b on lähinnä prosessin toimintaa varmistava vaihe mahdollista sen erottamisen, onko jakelupaketti keskeneräinen vai kokonaan olematon. Koska keskeneräinen jakelupaketti ei ole lainkaan noudettavissa, niin vaihe 4.d toiston kera aiheuttaa tarvittavan seurannan niin kauan, kunnes jakelupaketti saadaan noudettua.

## LIITE B. SIIRTOPAKETIN TARKASTUSRAPORTTI

Tarkastusraportti sisältää joukon PREMIS XML -muodon mukaisia objekteja, tapahtumia ja agentteja (tapahtuman suorittajia). Tässä liitteessä jokainen tällainen objekti, tapahtuma ja agentti on tiivistetysti avattu. Lueteltuja objekteja, tapahtumia ja agentteja on tarkastusraportissa kutakin yksi, ellei taulukon kohdalla ole muuta mainittu. Tapahtuman nimi ja lopputulos sekä tapahtumaan liittyvien objektien tiedostonimet näkyvät myös erillisessä HTML-muotoisessa ihmisluettavassa raportissa.

### Objektit

Vastaanotossa käytetyistä ja luoduista kokonaisuuksista muodostetaan PREMIS XML -muotoiset esitykset, joita kutsutaan PREMIS-objekteiksi. Eri objekteihin kohdistuu vastaanotossa erilaisia tapahtumia (ks. luku [Q](#)), joita suorittavat agentit (ks. luku [Q](#)). Näitä kokonaisuuksia ovat siirtopaketti, digitaalinen allekirjoitus, METS-dokumentti, digitaalinen objekti sekä luotu säilytyspaketti. Tässä kappaleessa on kuvattu näitä kutakin kokonaisuutta vastaavien PREMIS-objektien sisällöt.

Siirtopaketti:

Elementti	Määrittely
<object>	Siirtopaketti tarkastusraportissa
@xsi:type	"representation" (PREMIS-objektin tyyppi)
-<objectIdentifier>	Siirtopaketin tunniste tarkastusraportissa
--<objectIdentifierType>	"preservation-sip-id" (tunnisteen tyyppi)
--<objectIdentifierValue>	Tunnisteen arvo
-<originalName>	ZIP- tai TAR-tiedoston nimi, jossa siirtopaketti alkujaan sijaitti
-<environment>	-
--<dependency>	Riippuvuusuhde siirtopakettiin
---<dependencyIdentifier>	Siirtopaketin tunniste METS-dokumentissa
----<dependencyIdentifierType>	"mets:OBJID"
----<dependencyIdentifierValue>	Siirtopaketin tunnisteen arvo
--<dependency>	Riippuvuusuhde sopimustunnisteeseen
---<dependencyIdentifier>	Sopimustunniste METS-dokumentissa
----<dependencyIdentifierType>	preservation-contract-id
----<dependencyIdentifierValue>	Sopimustunnisteen arvo

Siirtopaketin digitaalinen allekirjoitus:

Elementti	Määrittely
<object>	Siirtopaketin digitaalinen allekirjoitustiedosto tarkastusraportissa
@xsi:type	"representation" (PREMIS-objektin tyyppi)
-<objectIdentifier>	Digitaalisen allekirjoitustiedoston tunniste tarkastusraportissa
--<objectIdentifierType>	"preservation-signature-id" (tunnisteen tyyppi)
--<objectIdentifierValue>	Tunnisteen arvo
-<originalName>	Digitaalisen allekirjoituksen tiedostonimi
-<relationship>	Suhde siirtopakettiin
--<relationshipType>	"structural" (rakenteellinen suhde)
--<relationshipSubType>	"is included in" (sisältyy)
--<relatedObjectIdentification>	Siirtopaketin tunniste tarkastusraportissa
---<relatedObjectIdentifierType>	"preservation-sip-id" (siirtopaketin tunnisteen tyyppi)
---<relatedObjectIdentifierValue>	Siirtopaketin tunnisteen arvo

Deleted: B.2

Deleted: B.3

## METS-dokumentti:

Elementti	Määrittely
<object>	METS-dokumentti tarkastusraportissa
@xsi:type	"representation" (PREMIS-objektin tyyppi)
-<objectIdentifier>	METS-dokumentin tunniste tarkastusraportissa
--<objectIdentifierType>	"preservation-mets-id" (tunnisteen tyyppi)
--<objectIdentifierValue>	Tunnisteen arvo
-<originalName>	METS-dokumentin tiedostonimi
-<relationship>	Suhde siirtopakettiin
--<relationshipType>	"structural" (rakenteellinen suhde)
--<relationshipSubType>	"is included in" (sisältyy)
---<relatedObjectIdentification>	Siirtopaketin tunniste tarkastusraportissa
---<relatedObjectIdentifierType>	"preservation-sip-id" (siirtopaketin tunnisteen tyyppi)
---<relatedObjectIdentifierValue>	Siirtopaketin tunnisteen arvo

Digitaalinen objekti: Jokaiselle digitaaliselle objektille tulee raporttiin oma PREMIS-objektinsa.

Elementti	Määrittely
<object>	Digitaalinen objekti tarkastusraportissa
@xsi:type	"representation" (PREMIS-objektin tyyppi)
-<objectIdentifier>	Digitaalisen objektin tunniste tarkastusraportissa
--<objectIdentifierType>	"preservation-object-id" (tunnisteen tyyppi)
--<objectIdentifierValue>	Tunnisteen arvo
-<originalName>	Digitaalisen objektin tiedostonimi
-<environment>	-
--<dependency>	Riippuvuussuhde digitaaliseen objektiin
---<dependencyIdentifier>	Digitaalisen objektin tunniste METS-dokumentissa
----<dependencyIdentifierType>	Digitaalisen objektin tunnisteen tyyppi
----<dependencyIdentifierValue>	Digitaalisen objektin tunnisteen arvo
-<relationship>	Suhde siirtopakettiin
--<relationshipType>	"structural" (rakenteellinen suhde)
--<relationshipSubType>	"is included in" (sisältyy)
---<relatedObjectIdentification>	Siirtopaketin tunniste tarkastusraportissa
---<relatedObjectIdentifierType>	"preservation-sip-id" (siirtopaketin tunnisteen tyyppi)
---<relatedObjectIdentifierValue>	Siirtopaketin tunnisteen arvo

## Säilytyspaketti:

Elementti	Määrittely
<object>	Säilytyspaketti tarkastusraportissa
@xsi:type	"representation" (PREMIS-objektin tyyppi)
-<objectIdentifier>	Säilytyspaketin tunniste tarkastusraportissa
--<objectIdentifierType>	"preservation-aip-id" (tunnisteen tyyppi)
--<objectIdentifierValue>	Tunnisteen arvo
-<originalName>	Säilytyspaketin tiedostonimi
-<relationship>	Suhde siirtopakettiin
--<relationshipType>	"derivation" (johdannainen)
--<relationshipSubType>	"has source" (on olemassa lähde)
---<relatedObjectIdentification>	Siirtopaketin tunniste tarkastusraportissa
---<relatedObjectIdentifierType>	"preservation-sip-id" (siirtopaketin tunnisteen tyyppi)
---<relatedObjectIdentifierValue>	Siirtopaketin tunnisteen arvo



## Tapahtumat

Vastaanotossa suoritetuista toimenpiteistä muodostetaan PREMIS XML -muotoiset esitykset, joita kutsutaan PREMIS-tapahtumiksi. Vastaanoton aikana eri toimenpiteet kohdistuvat eri objekteihin (ks. luku [Q](#)), ja niitä suorittavat agentit (ks. luku [Q](#)). Näitä toimenpiteitä on useita, ja tässä kappaleessa on kuvattu kutakin toimenpidettä esittelevien PREMIS-tapahtumien sisällöt. Tapahtumalohko tietynlaiselle toimenpiteelle (esim. METS-lisävalidointi tai digitaalisen objektin validointi) voi esiintyä raportissa useita kertoja. Seuraavana on esitetty taulukkona tapahtumarunko, joka on sama kaikille tapahtumille. Tämän jälkeen on luettelona esitetty vastaanotossa tavallisesti esiintyvät tapahtumat sekä tarkennuksia tapahtumarungon muutamille elementeille.

Tapahtumarunko:

Elementti	Määrittely
<event>	-
--<eventIdentifier>	Tapahtuman tunniste
--<eventIdentifierType>	"preservation-event-id" (tunnisteen tyyppi)
--<eventIdentifierValue>	Tunnisteen arvo
-<eventType>	Tapahtuman tyyppi, ks. tarkennus tapahtumista
-<eventDateTime>	Tapahtuman aikaleima
-<eventDetail>	Tapahtuman nimi, ks. tarkennus tapahtumista
-<eventOutcomeInformation>	-
--<eventOutcome>	Tapahtuman lopputulos: success tai failure
--<eventOutcomeDetail>	-
---<eventOutcomeDetailNote>	Tekstimuotoinen lisätieto tapahtuman lopputuloksesta
---<eventOutcomeDetailExtension>	XML-muotoinen lisätieto tapahtuman lopputuloksesta
-<linkingAgentIdentifier>	Viittaus agenttiin tarkastusraportissa, ks. tarkennus tapahtumista
-<linkingAgentIdentifierType>	Agentin tunnisteen tyyppi
-<linkingAgentIdentifierValue>	Agentin tunnisteen arvo
-<linkingObjectIdentifier>	Viittaus objektiin tarkastusraportissa, ks. tarkennus tapahtumista
-<linkingObjectIdentifierType>	Objektin tunnisteen tyyppi
-<linkingObjectIdentifierValue>	Objektin tunnisteen arvo

Siirtopaketin siirto:

Elementti	Määrittely
-<eventType>	"transfer"
-<eventDetail>	"Transfer of submission information package"
-<linkingAgentIdentifier>	Viittaus aineiston siirtäneeseen käyttäjätunnukseen
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

Siirtopaketin purku:

Elementti	Määrittely
-<eventType>	"unpacking"
-<eventDetail>	"Unpacking of submission information package"
-<linkingAgentIdentifier>	Viittaus purkusovellukseen
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

Virustarkastus:

Elementti	Määrittely
-<eventType>	"virus check"
-<eventDetail>	"Virus check of transferred files"
-<linkingAgentIdentifier>	Viittaus tarkastusovellukseen

Deleted: B.1

Deleted: B.3

Elementti	Määrittely
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

Digitaalisen allekirjoituksen tarkastus:

Elementti	Määrittely
-<eventType>	"digital signature validation"
-<eventDetail>	"Submission information package digital signature validation"
-<linkingAgentIdentifier>	Viittaus tarkastussovellukseen
-<linkingObjectIdentifier>	Viittaus digitaaliseen allekirjoitukseen

Sopimuksen tarkastus:

Elementti	Määrittely
-<eventType>	"validation"
-<eventDetail>	"Validation of service contract properties"
-<linkingAgentIdentifier>	Viittaus tarkastussovellukseen
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

METS-skeemavalidointi:

Elementti	Määrittely
-<eventType>	"validation"
-<eventDetail>	"METS schema validation"
-<linkingAgentIdentifier>	Viittaus validointisovellukseen
-<linkingObjectIdentifier>	Viittaus METS-dokumenttiin

METS-lisävalidointi:

Elementti	Määrittely
-<eventType>	"validation"
-<eventDetail>	"Additional METS validation of required features"
-<linkingAgentIdentifier>	Viittaus validointisovellukseen
-<linkingObjectIdentifier>	Viittaus METS-dokumenttiin

Digitaalisten objektien eheystarkastus: tapahtuma kattaa kaikkien siirtopaketissa olleiden digitaalisten objektien eheystarkastuksen.

Elementti	Määrittely
-<eventType>	"fixity check"
-<eventDetail>	"Fixity check of digital objects in submission information package"
-<linkingAgentIdentifier>	Viittaus tarkastussovellukseen
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

Digitaalisten objektien validointi: Jokaiselle digitaaliselle objektille tulee vähintään yksi tapahtuma.

Elementti	Määrittely
-<eventType>	"validation"
-<eventDetail>	"Digital object validation"
-<linkingAgentIdentifier>	Viittaus validointisovellukseen
-<linkingObjectIdentifier>	Viittaus digitaaliseen objektiin

Siirtopaketin validoinnin yhteenveto:

Elementti	Määrittely
-<eventType>	"validation"
-<eventDetail>	"Validation compilation of submission information package"

Elementti	Määrittely
-<linkingAgentIdentifier>	Viittaus validointisovellukseen
-<linkingObjectIdentifier>	Viittaus siirtopakettiin

Säilytyspaketin luominen:

Elementti	Määrittely
-<eventType>	"information package creation"
-<eventDetail>	"Creation of archival information package"
-<linkingAgentIdentifier>	Viittaus luontisovellukseen
-<linkingObjectIdentifier>	Viittaus luotuun säilytyspakettiin

Säilytysvastuun siirtyminen:

Elementti	Määrittely
-<eventType>	"accession"
-<eventDetail>	"Preservation responsibility change to the digital preservation service"
-<linkingAgentIdentifier>	Viittaus siirtosovellukseen
-<linkingObjectIdentifier>	Viittaus säilytyspakettiin

### Agentit

Vastaanotossa suoritettujen toimenpiteiden (ks. luku [9](#)) suorittajista muodostetaan PREMIS XML -muotoiset esitykset, joita kutsutaan PREMIS-agenteiksi. Suorittajana on joko hyödyntävä organisaatio tai ohjelma. Tässä kappaleessa on esitelty näitä suorittajia kuvaavien PREMIS-agenttien sisällöt.

Aineiston siirtäneen hyödyntävän organisaation käyttäjätunnus:

Elementti	Määrittely
<agent>	-
-<agentIdentifier>	Agentin tunniste
--<agentIdentifierType>	Tunnisteen tyyppi
--<agentIdentifierValue>	Tunnisteen arvo
-<agentName>	Hyödyntävän organisaation käyttäjätunnus
-<agentType>	"organization" (agentin tyyppi)

PAS-ohjelmistokomponentti: jokaiselle raportissa viitatuille ohjelmistokomponentille tulee oma agenttinsa.

Elementti	Määrittely
<agent>	-
-<agentIdentifier>	Agentin tunniste
--<agentIdentifierType>	Tunnisteen tyyppi
--<agentIdentifierValue>	Tunnisteen arvo
-<agentName>	Ohjelmakoodin nimi
-<agentType>	"software" (agentin tyyppi)

Deleted: B.2

## LÄHTEET

- [FI\_FORMATS] Kansalliset PAS-palvelut. Säilytys- ja siirtokelpoiset tiedostomuodot. [urn:nbn:fi-fe2020100578095](https://nbn.fi/urn:nbn:fi-fe2020100578095)
- [FI\_META] Kansalliset PAS-palvelut. Aineistojen ja niiden metatietojen paketointi pitkäaikais säilytykseen. [urn:nbn:fi-fe2020100578093](https://nbn.fi/urn:nbn:fi-fe2020100578093)
- [IETF\_SFTP] Internet Engineering Task Force Internet-Draft: SSH File Transfer Protocol, J. Galbraith, O. Saarenmaa, July 2006. <https://tools.ietf.org/html/draft-ietf-secsh-filexfer-13>
- [INVENIO] Invenio. <https://invenio-software.org/>
- [JSEND] OmniTI Labs: JSend. <https://github.com/omniti-labs/jsend>
- [JSON\_API] JSON API. <https://jsonapi.org/>
- [LUCENE] Apache Lucene - Query Parser Syntax. [https://lucene.apache.org/core/3\\_6\\_0/queryparsersyntax.html](https://lucene.apache.org/core/3_6_0/queryparsersyntax.html)
- [LUQUM] Luqum's Documentation. <https://luqum.readthedocs.io/>
- [PREMIS] Data Dictionary for Preservation Metadata: PREMIS version 2.2, July 2012. <https://www.loc.gov/standards/premis/v2/premis-2-2.pdf>
- [RFC\_7159] Internet Engineering Task Force Request for Comments: The JavaScript Object Notation (JSON) Data Interchange Format. <https://tools.ietf.org/html/rfc7159>
- [RFC\_7230] Internet Engineering Task Force Request for Comments: Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing. <https://tools.ietf.org/html/rfc7230>
- [RFC\_7231] Internet Engineering Task Force Request for Comments: Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content. <https://tools.ietf.org/html/rfc7231>
- [TUS] TUS.io Open Protocol for Resumable File Uploads. <https://tus.io/>

Field Code Changed

Field Code Changed

Field Code Changed