



# Mikä Kulttuuriperintö-PAS-palvelu on?

## Digitaalisten aineistojen säilyttäminen Kulttuuriperintö-PAS-palvelussa

4.10.2022

Juha Lehtonen



# Säilyttämisen palvelut

- CSC tuottaa OKM:lle keskitettyjä PAS-palveluita
  - Kulttuuriperinnön säilyttämiseen (tuotannossa vuodesta 2015)
  - Tutkimukseen liittyvien aineistojen säilyttämiseen (tuotannossa vuodesta 2019)
- Aineistojen omistajuus säilyy **aina** hyödyntävällä organisaatiolla
  - Vain hyödyntävällä organisaatiolla on pääsy säilytettäviin aineistoihin
- Palvelut tuotetaan OKM:n rahoituksella ja ovat maksuttomia niitä hyödyntävälle organisaatiolle
  - Palvelut saa käyttöönsä OKM:n päätöksellä
- Kulttuuriperintö-PAS-palvelussa sopimuksia 15
  - Kansallisarkisto, Kansalliskirjasto, Kavi, Kansallisgalleria, Museovirasto, Celia, Kotus, Musiikkiarkisto, SLS, Tietoarkisto
- Fairdata PAS-palvelussa sopimuksia 5
  - Turun yliopisto, Itä-Suomen yliopisto, Helsingin yliopisto, GTK, Tampereen yliopisto

Kulttuuriperintö-  
PAS-palvelu 

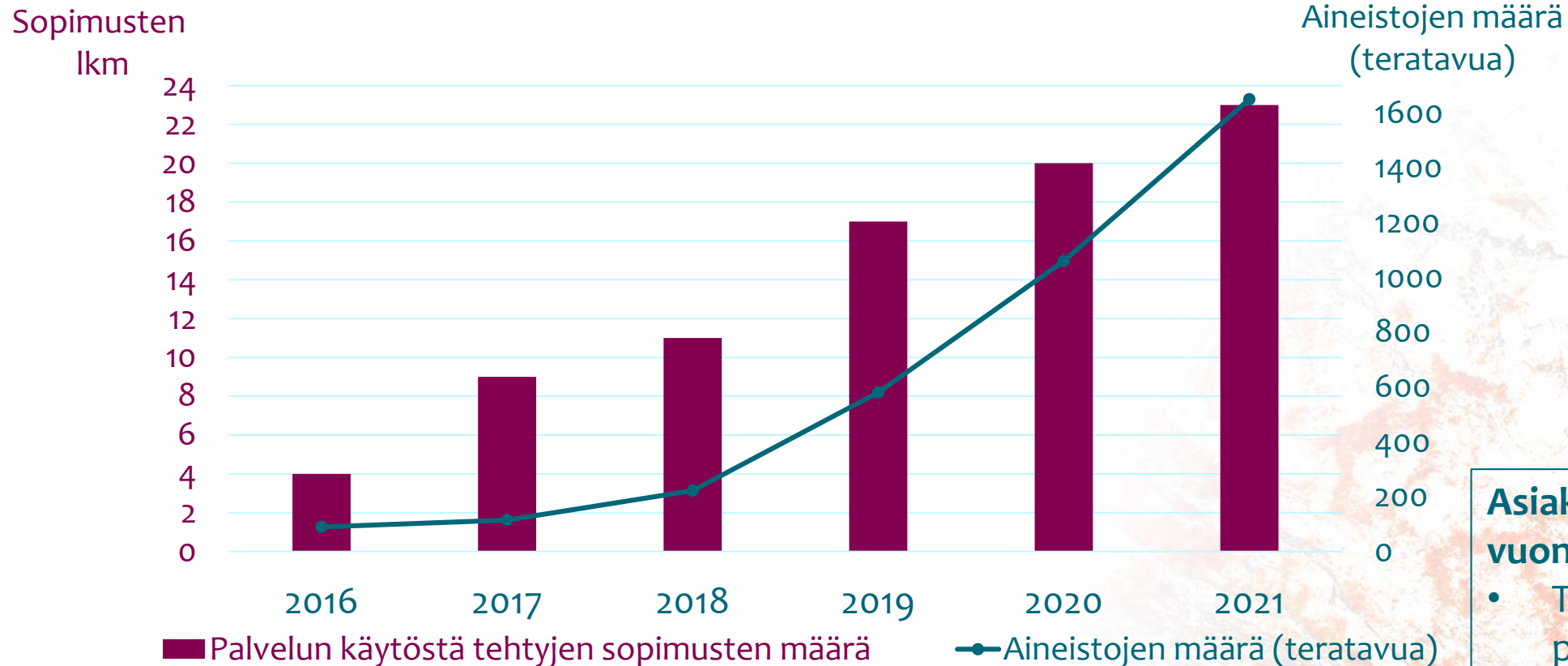
Fairdata  
PAS-palvelu 

Yhteinen PAS-ratkaisu





# Pitkäaikaissäilytyksen palvelujen käytön kehittyminen 2016–2021



## Asiakastyytyväisyys oli vuonna 2021 varsin hyvä:

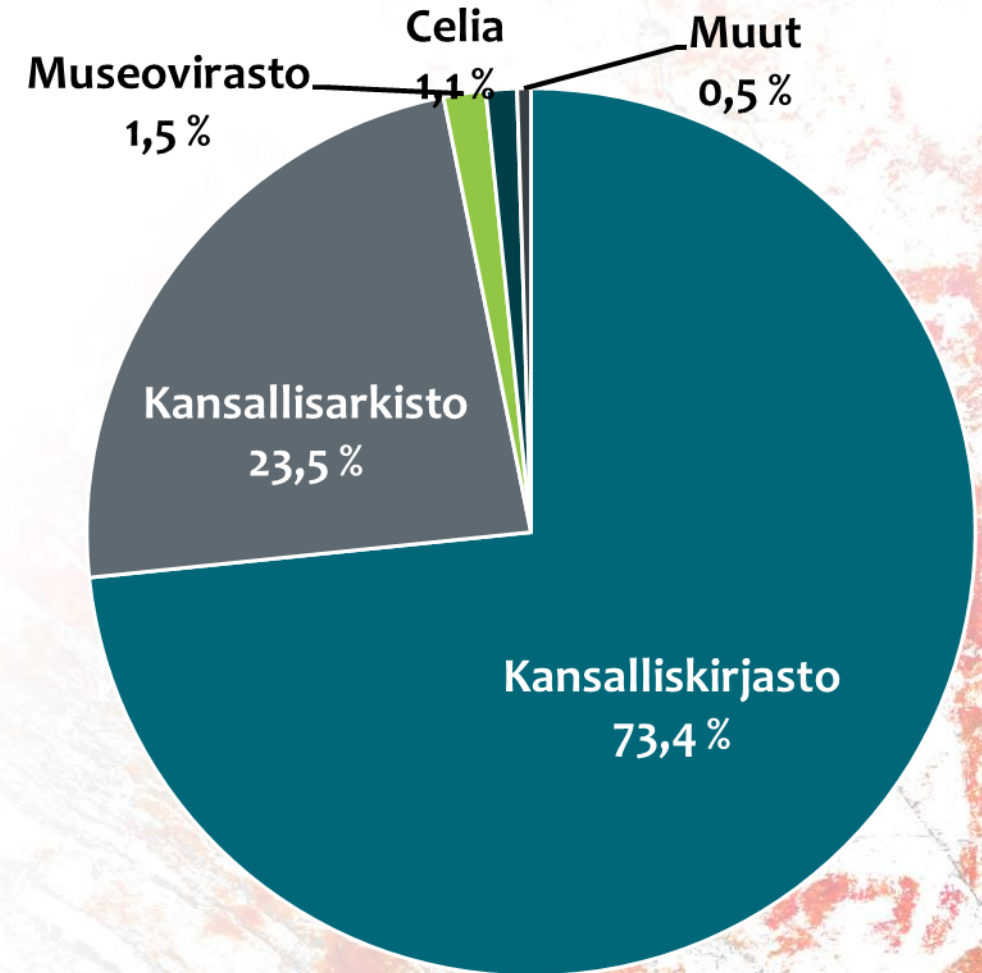
- Tyytyväisyys PAS-palveluihin: ka. 5,5/6
- Tyytyväisyys asiakaspalveluun ja neuvontaan: ka. 5,0/6

Huom. Mukana ovat Kulttuuriperintö-PAS-palvelu (tuotannossa 11/2015 alkaen) ja Fairdata PAS -palvelu (tuotannossa 12/2019). Vuositilastot osoittavat vuoden lopun tilanteen.

Lähde: PAS-palveluiden käyttötilastot.

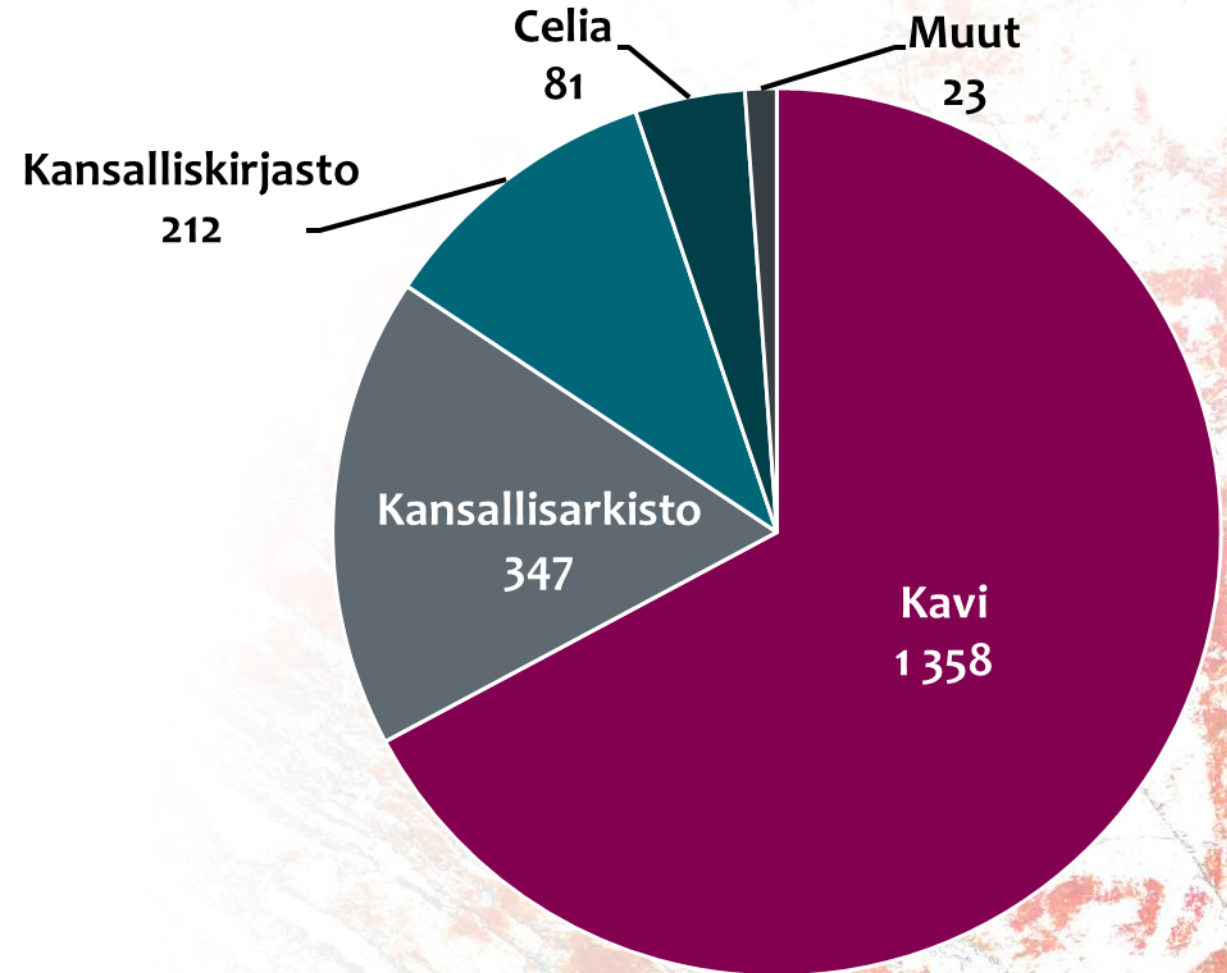
# Säilytyksessä yli 2,6 miljoonaa tietopakettia

- Säilytykseen on hyväksytty yli 2,6 miljoonaa pakettia (syyskuu 2022)
- Sisältää digitaaliset objektit sekä kuvailevaa, hallinnollista ja rakenteellista metatietoa
- Säilyttämisen toimenpiteet kohdentuvat erityisesti tietopaketteihin
  - Laaduntarkastus, virkistäminen, migraatio, päivittäminen, ...
- Tietopakettien rakenne määritellään yhdessä hyödyntävien organisaatioiden kanssa



# Aineistomäärä ylittänyt 2 petatavua

- Säilytykseen hyväksytyjen aineistojen määrä on jo yli 2 petatavua (syyskuu 2022)
- Säilytysstrategian mukaisesti aineistosta on useampi kopio useammassa sijainnissa
  - Todellinen bittien määrä moninkertainen
- Viereisessä kaaviossa organisaatioiden aineistomäärät on esitetty teratavuissa





# Kulttuuriperintö-PAS-palvelun käyttötarkoituksia (1/2)

Organisaatio	Käyttötarkoitus
Celia	Master-arkisto ja pitkäaikaisesti säilytettävät valitut uudet äänikirjat
Kansallinen audiovisuaalinen instituutti	Valikoitu osa kotimaisen elokuvan digitoitavista aineistoista
Kansallisarkisto	Kansallisarkiston vastaanottamat alkujaan digitaaliset valtionhallinnon asiakirjalliset aineistot
Kansallisarkisto	VAPA-järjestelmään siirretyt tietoaineistot
Kansallisarkisto	Kansallisarkiston massadigitointi-hankkeen aineistot
Kansallisarkisto	Kansallisarkiston digitaaliarkistosta siirrettävät aineistot ja takautuvan digitoinnin aineistot
Kansallisarkisto	Kansallisarkiston yksinomaan digitaalisessa muodossa olevat yksityisarkistoaineistot
Kansallisgalleria	Kiasman mediataiteen teosten pitkäaikaissäilytys

# Kulttuuriperintö-PAS-palvelun käyttötarkoituksia (2/2)

Organisaatio	Käyttötarkoitus
Kansalliskirjasto	Kansalliskirjaston digitoimat kulttuuriperintöaineistot
Kansalliskirjasto	Kulttuuriaineistolain nojalla kerätyt aineistot
Kotimaisten kielten keskus Kotus	Kotuksen kielentutkimus- ja kulttuuriperintöaineistojen pitkäaikaissäilytys
Museovirasto	Kulttuuriympäristön tutkimusraportit
Musiikkiarkisto	Musiikkiarkiston pitkäaikaissäilytettävät aineistot
Svenska Litteratursällskapet SLS	SLS:n pitkäaikaissäilytettävät aineistot
Yhteiskuntatieteellinen tietoarkisto	Tietoarkiston arkistoimien tutkimusaineistojen kokoelman pitkäaikaissäilytys

## Mitä hyödyntäminen edellyttää organisaatiolta?

- Sopimusmalliin tutustumista
- PAS-palvelun hyödyntämisestä sopimista OKM:n kanssa
- Säilytettävän aineistokokonaisuuden tunnistamista
- Käyttöönottoprosessin läpikäymistä yhdessä CSC:n kanssa
- Aineistojen riittävää kuvailua
- Aineistojen paketointia PAS-määritysten mukaisesti
- Aineistojen siirtämistä PAS-palveluun

**Kaikki tämä edellyttää hyödyntävältä organisaatiolta henkilöresursseja palvelun käyttöönottoon ja käyttöön**

mutta...

**PAS-palvelu  
tukee  
organisaatioita  
kaikissa  
vaiheissa**



# PAS-palvelun hyödyntäminen on yhteistyötä



PAS ei ole “itsepalvelu”-palvelu

Ratkaistaan ongelmat yhdessä.  
Kysyvä ei tieltä eksy.



PAS-palvelu tukee organisaatioita  
kaikissa vaiheissa

Tutustumisesta käyttöönottoon ja  
käytön aikana.



Keskustellaan parhaat tavat  
alusta asti yhdessä

Ja jatketaan yhteydenpitoa.

# Kysely: Museoiden valmius ja kiinnostus liittyä PAS-palvelun hyödyntäjiksi



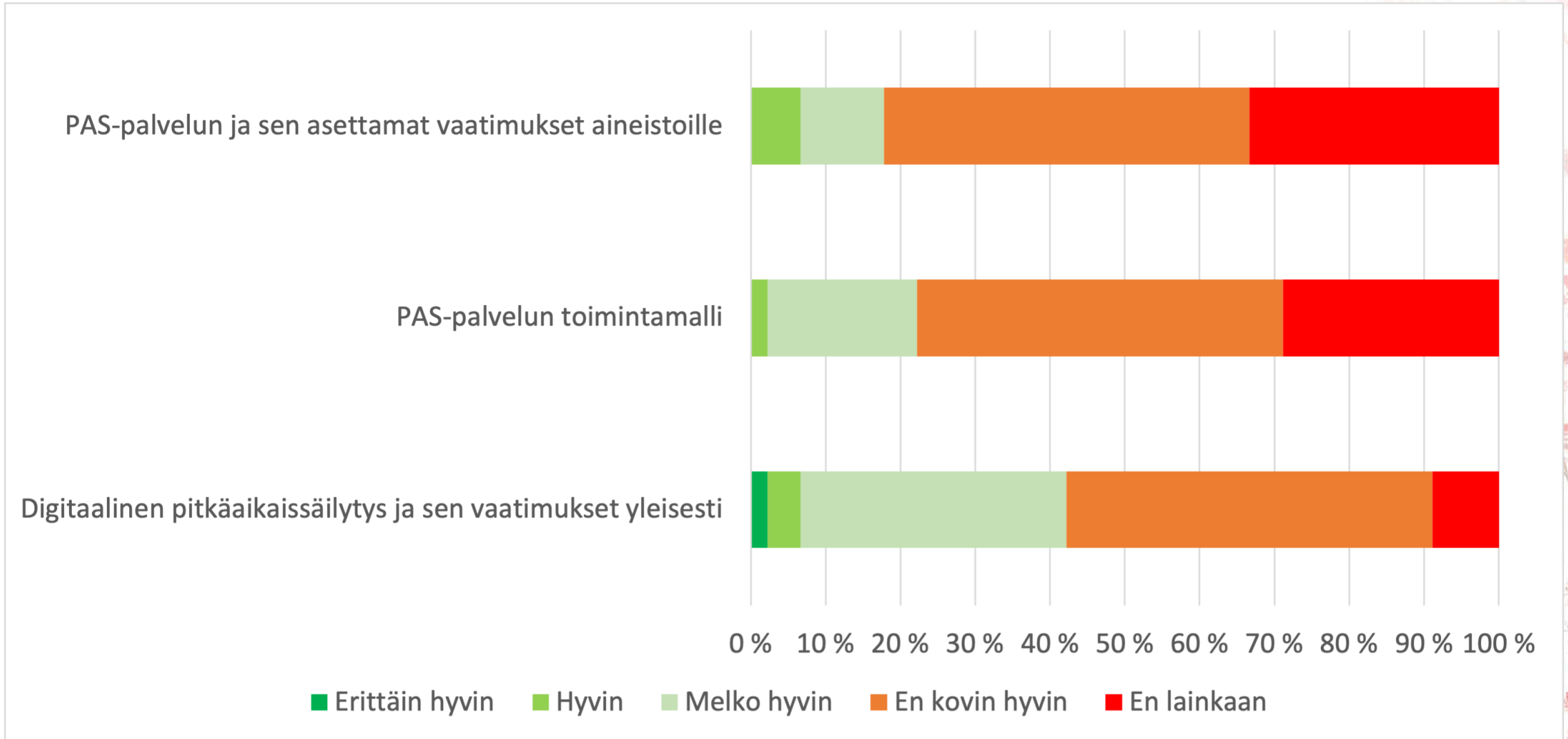
- OKM:n ja CSC:n toteuttama kysely joulukuussa 2021
- Ammatillisesti hoidetuille museoille suunnattu
  - Erityisesti museoille, jotka eivät vielä hyödynnä PAS-palvelua
- PAS-valmiuden ja -kiinnostuksen kartoitus
- PAS-palveluiden tunnettavuuden kartoitus

# Kysely: Vastaajat

- Kysely kohdennettiin noin 150 museonjohtajalle, joista vastasi 30%
- Kyselyyn vastasi 45 museota
  - 23 oli alueellisia tai valtakunnallisia vastuumuseoita
  - 22 muita ammatillisia museoita
- Kyselyn vastaukset eivät ole sitovia
- On mahdollisuus osallistua PAS-toimintaan, vaikka ei olisi vastannut kyselyyn

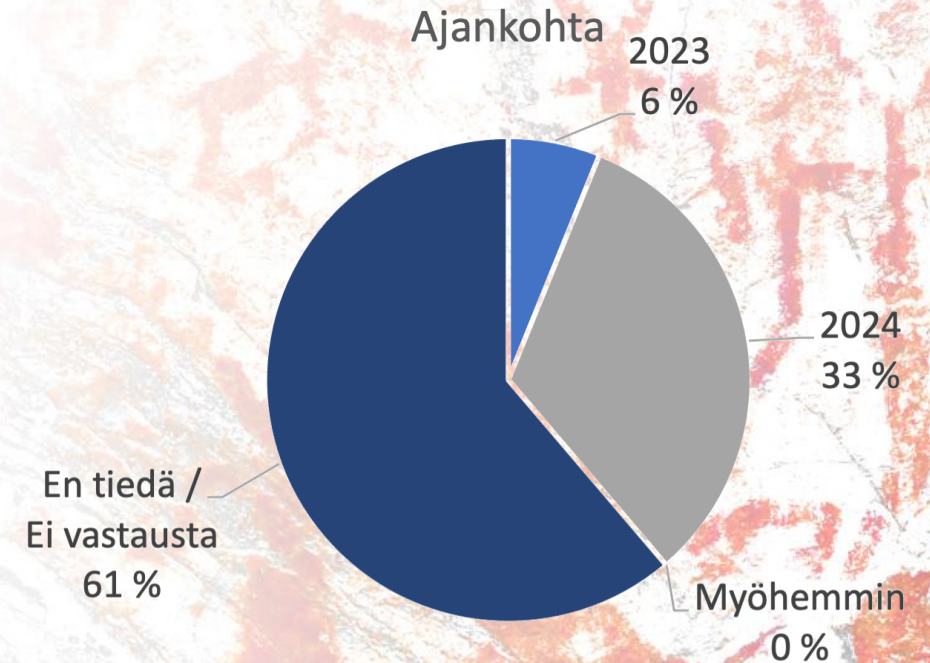
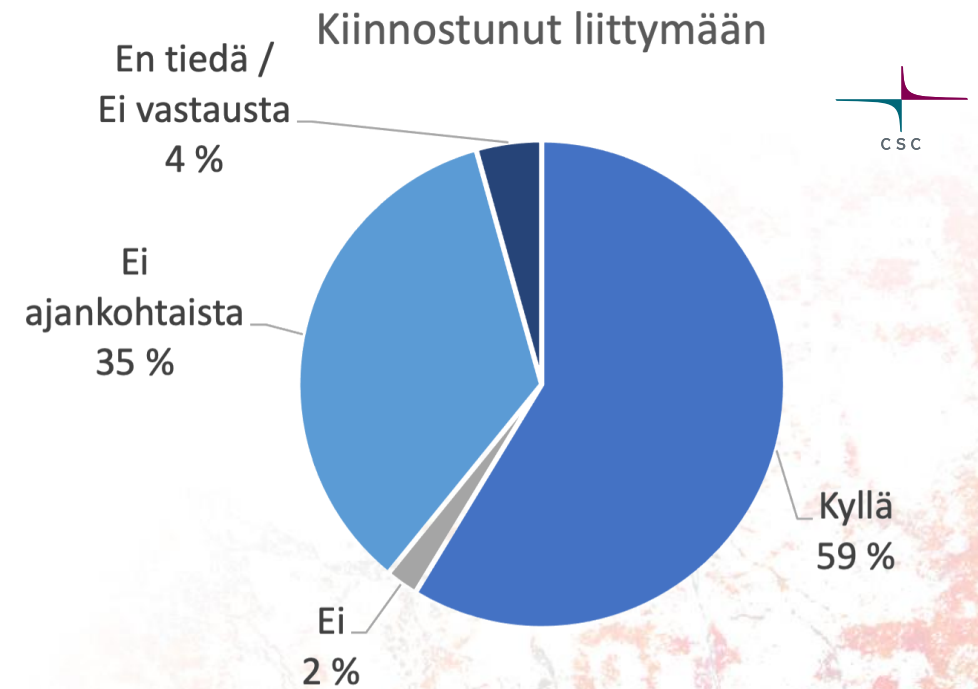


# Kysely: "Miten hyvin tunnette...?"

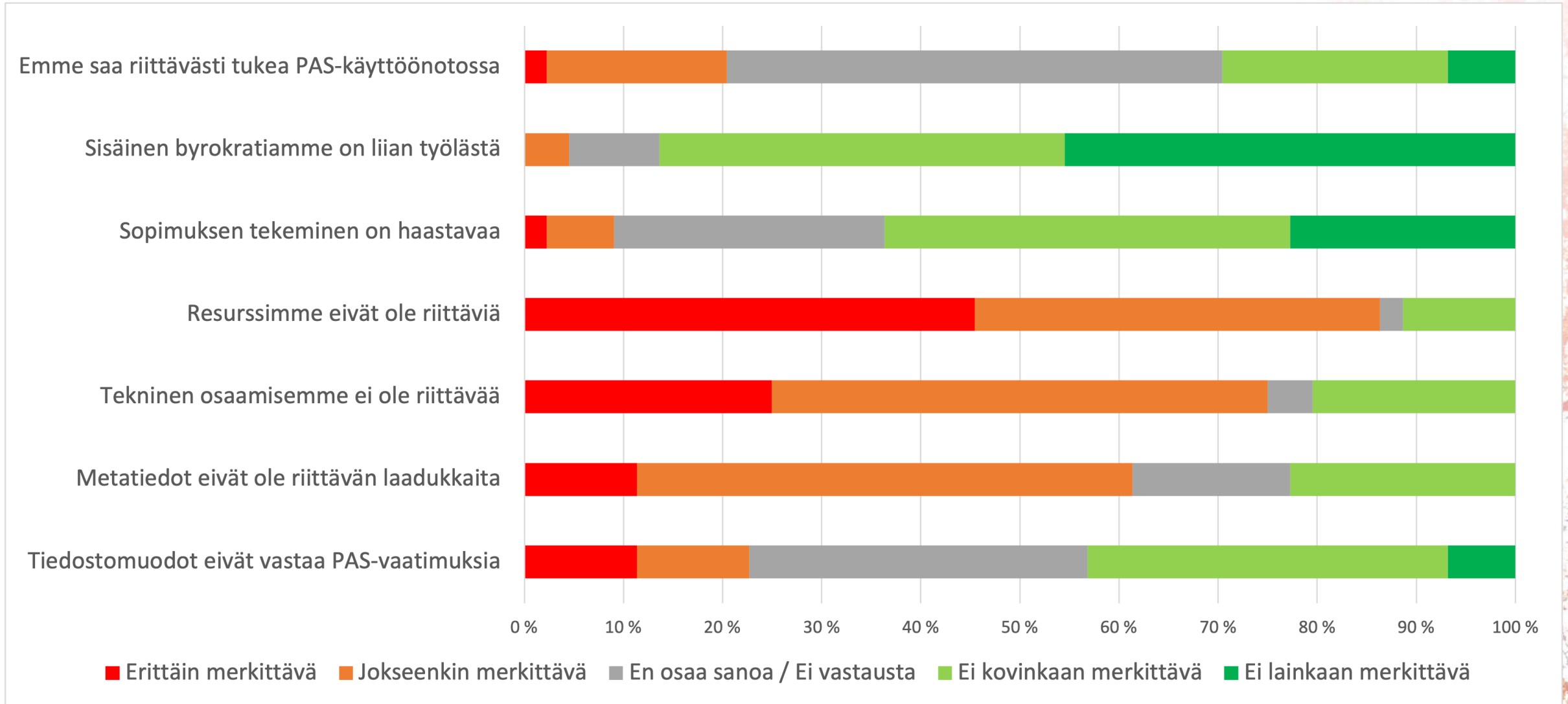


# Kysely: Liittyminen PAS-palvelun hyödyntäjäksi

- Vastanneista noin 3/5 on kiinnostunut liittymään PAS-palvelun hyödyntäjäksi
  - Lähes kaikki muut vastaajat eivät kokeneet liittymistä ajankohtaiseksi aivan vielä.
- Vastanneista noin 2/5 arvioi liittymisen ajankohtaiseksi joko vuonna 2023 tai 2024
  - Muut vastaajat eivät joko osanneet tai halunneet määrittellä ajankohtaa



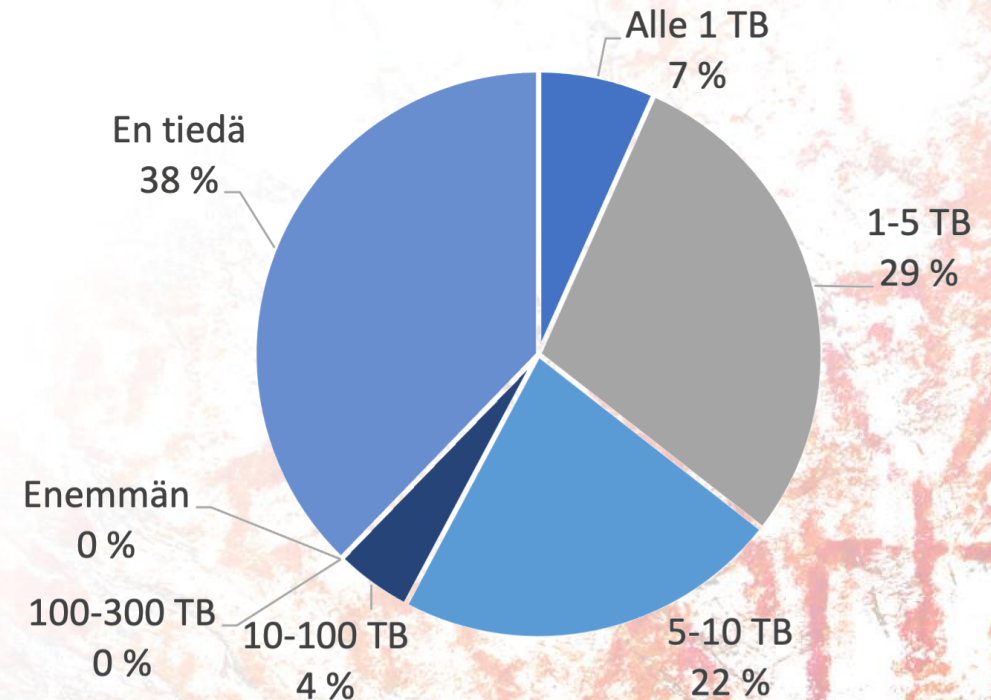
# Kysely: Haasteet PAS-palvelun käyttöönottamiseksi



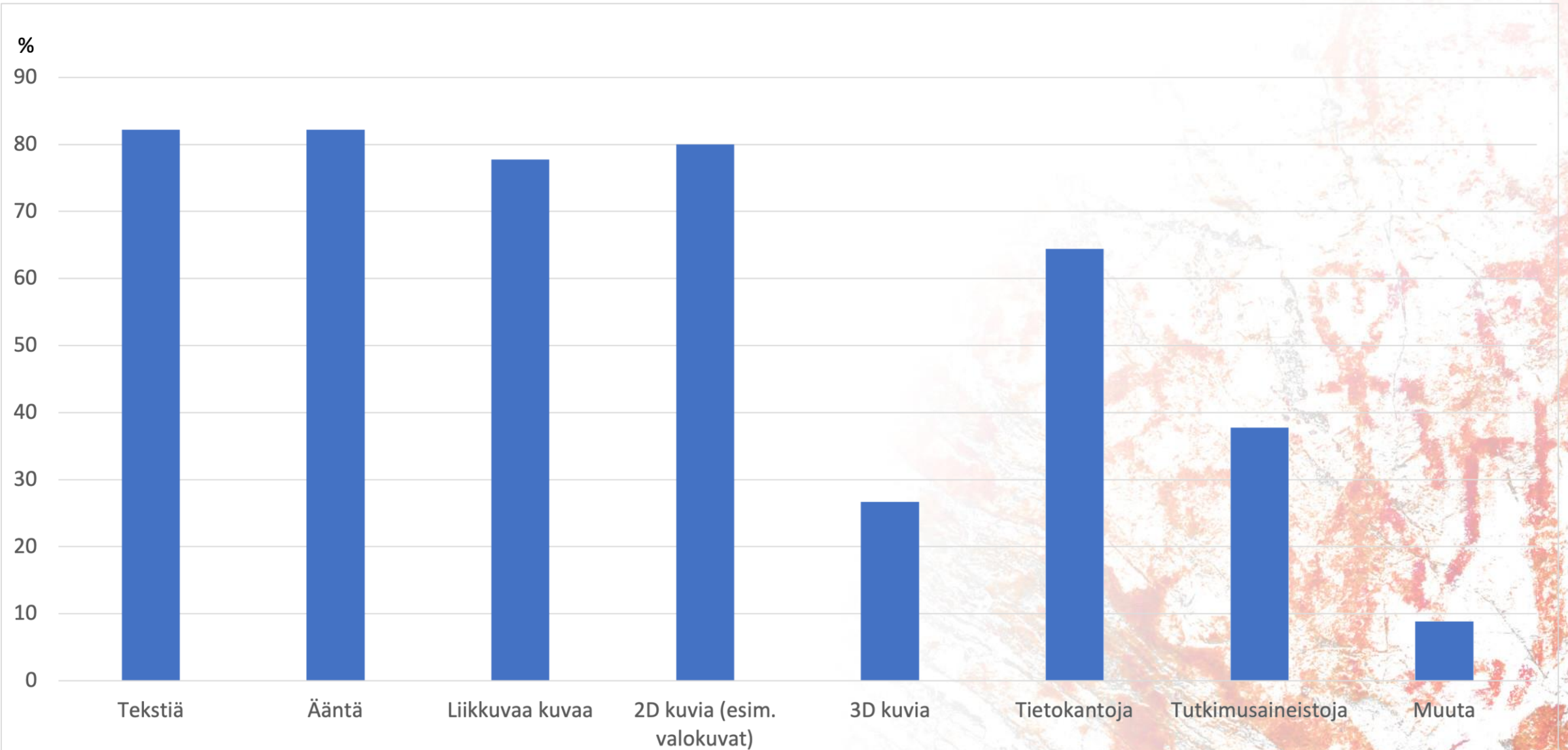


# Kysely: Aineistomäärä PAS-palveluun

- Yli puolet vastaajista: alle 10 teratavua
- Keskiarvo 13 teratavua (arvioineiden kesken)
  - Toisilla aineistoa on alle yksi teratavu ja toisilla 100 teratavua
- Yli kolmannes ei osannut tai halunnut arvioida aineistomäärää
- Vuosina 2023–2024 mukaan haluavien museoiden arvioitu aineistomäärä on noin 240 teratavua
  - PAS-palvelun näkökulmasta varsin maltillisesti



# Kysely: Aineistotyytit



# Kysely: Seuraavaksi

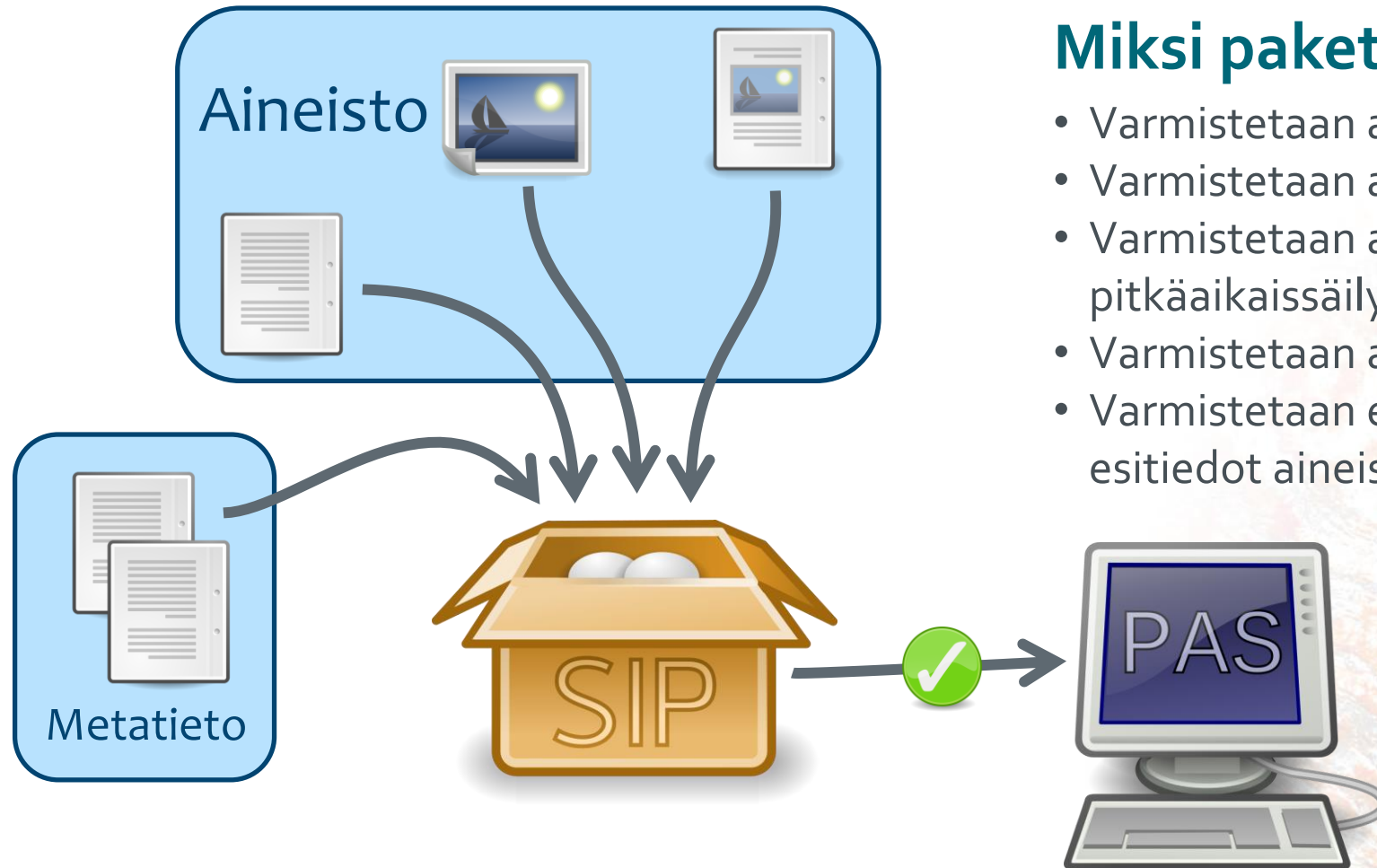
- OKM ja CSC tulevat käyttämään tämän kyselyn tuloksia pohjana, kun suunnitellaan PAS-palvelun käytön laajentamista erityisesti museosektorilla
- Kyselyn perusteella CSC parantaa museoiden valmiutta ja tietämystä PAS-palvelusta yleisesti tapahtumien ja muiden viestintäkanavien avulla
- Vuoden 2023 aikana museoiden tietämystä syvennetään konkreettisemmin teknisellä tasolla
  - Tavoitteena on, että viimeistään vuoden 2023 loppupuoliskosta alkaen useampi museo voisi liittyä PAS-palvelun hyödyntäjäksi
- Varsinainen vaiheistaminen liittymisestä sovitaan myöhemmin erikseen



# PAS-näkökulma mukaan jo alussa

- Omistajuus ja oikeudet, roolit ja vastuut
- Vaatimusten ja suositusten huomiointi
- Standardien noudattaminen (tiedostomuodot, metatietomuodot)
- Metatiedot (kuvailutiedot, tekninen, tapahtumahistoria...)
- Tiedostojen järjestäminen ja nimeämiskäytännöt
- Versiointi, varmuuskopiointi, tarkistussummien luonti
- Aineiston elinkaari. Mikä osuus menisi PAS-palveluun?
- Aineiston jatkokäyttö ja julkaisu
  - PAS-palvelu ei sovellu aineiston jokapäiväiseen käyttöön

# Aineistojen paketointi pitkäaikaissäilytykseen



## Miksi paketoidaan?

- Varmistetaan aineiston eheys
- Varmistetaan aineiston aitous
- Varmistetaan aineiston sopivuus pitkäaikaissäilytykseen
- Varmistetaan aineiston ymmärrettävyys
- Varmistetaan että PAS-palvelulla on riittävät esitiedot aineiston säilyttämiseksi

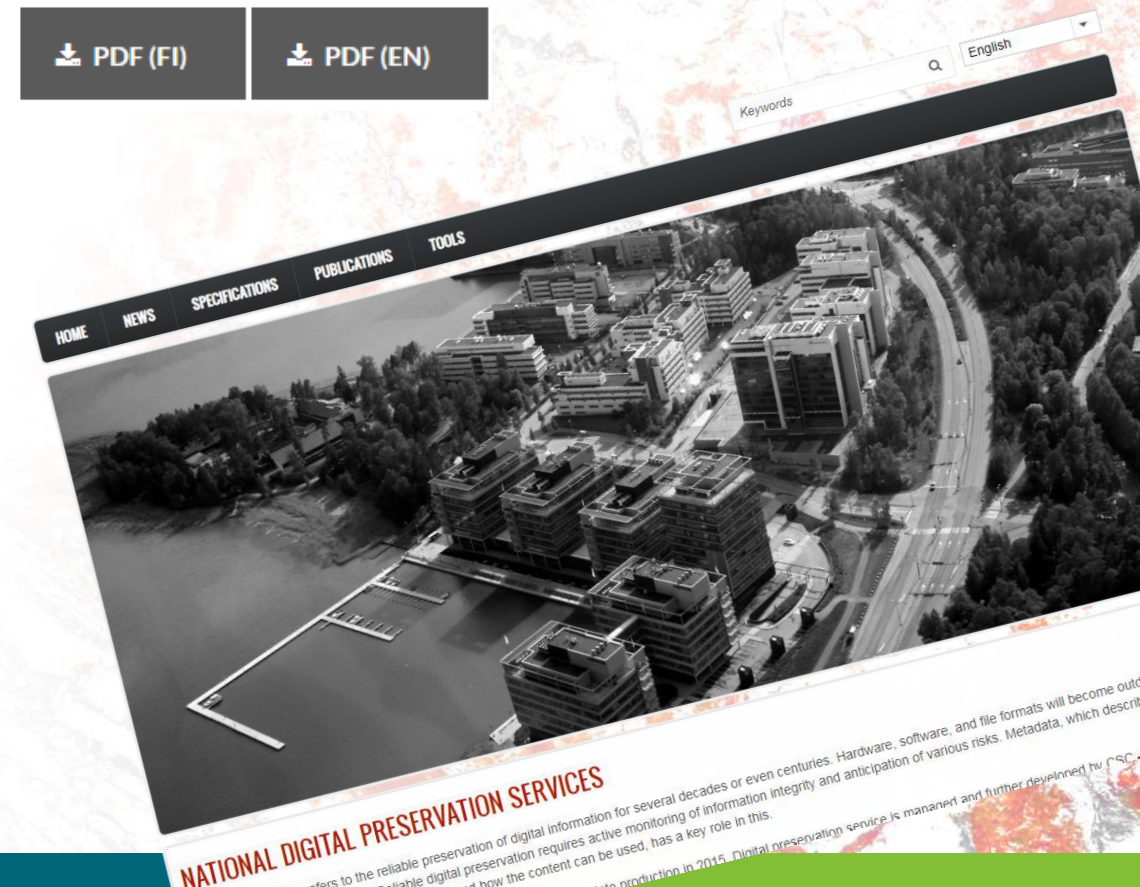
Aineistoja ei kannata paketoida etukäteen “varastoon”.

# Määrittelyt saatavilla

Myös englanniksi

- Aineistojen ja niiden metatietojen paketointi pitkäaikaissäilytykseen
- Säilytys- ja siirtokelpoiset tiedostomuodot
- PAS-palveluiden rajapinnat

<http://digitalpreservation.fi/specifications>





# Määrittelyistä käytäntöön



Aineistojen ja niiden metatietojen paketointi pitkäaikais... 159.38%

PAS-metatiedot-ja-aineiston-paketointi-1.7.3.pdf

AINEISTOJEN JA NIIDEN METATIETOJEN PAKETOINTI PITKÄAIKAISSÄILYTYKSEEN – 1.7.3

<b>SISÄLTÖ</b>	
<b>1 JOHDANTO</b>	4
1.1 PAS-palvelut	4
1.2 Aineiston kuvailu	4
1.3 Päivittyvä määrittely	5
<b>2 TIETOPAKETIN SISÄLLÖN KUVAILU</b>	8
2.1 METS-profiili	8
2.2 Skema	9
2.3 Linkit	9
2.4 Pitkäaikaissäilytyksen pakolliset metatiedot	9
2.4.1 Tunnisteet	10
2.4.1.1 Tunniste paketille	10
2.4.1.2 Tunniste aineistolle	11
2.4.1.3 Tunniste sopimukselle	11
2.4.1.4 Tunniste tiedostolle	11
2.4.1.5 Tunniste metatiedoille	11
2.4.2 Aikaleimat	12
2.4.2.1 Paketin luonti- tai muokkauksen metatietueiden luontiaika	12
2.4.2.2 Digitaalisten objektien luontiaika	13
2.4.3 Kuvailuformaatti	13
2.4.4 Tekniset metatiedot	14
2.4.4.1 Tiedostomuoto ja sen versio	14
2.4.4.2 Eheydet ja sen algoritmi	14
2.4.4.3 Tekniset piirteet	14
2.4.5 Rajoitukset	15
2.4.5.1 Aineistojen käyttörajoitukset	15
2.4.5.2 Säilyttämiseksi asetettavat rajoitukset	16
2.4.6 Aineiston synty- ja tapahtumahistoria	17
2.4.7 Rakennekartta	17
<b>3 SIIRTOPAKETIN TEKNINEN RAKENNE</b>	19
3.1 Siirtopaketin sisältö	19
3.2 Digitaalinen allekirjoitus	21
3.3 Metatietostandardit ja niiden versiot	21
3.4 Aineiston päivittäminen	22
<b>4 JAKELUPAKETIN TEKNINEN RAKENNE</b>	23
<b>LIITE A. KANSALLISEN METATIETOJEN PROFIILIT</b>	24
A.1 <meta>-elementti	24
A.2 <meta:generator>-elementti	25
A.3 <dmdSec>-elementti	26
A.4 <techMD>-elementti	27
A.5 <rightsMD>-elementti	28
A.6 <sourceMD>-elementti	29

Aineistojen ja niiden metatietojen paketointi pitkäaikaissäilytykseen | Digitalpreservation.fi

https://www.digitalpreservation.fi/specifications/metadata

Keywords

## AINEISTOJEN JA NIIDEN METATIETOJEN PAKETOINTI PITKÄAIKAISSÄILYTYKSEEN

Tämä dokumentti kuvaa kansallisissa pitkäaikaissäilytyspalvelussa (PAS-palvelussa) käytettävien METS-profiilien osalta.

**Versio**  
1.7.3

**Julkaisupäivämäärä**  
28.2.2021

**Tunniste**  
URN:urn:nbn:fi:20100578093

**Lisenssi**  
Creative Commons Attribution CC-BY-NC-SA 4.0

PDF HISTORIA

**di**

CSC – Tieteen tietotekniikan keskus Oy  
pas-support@csc.fi

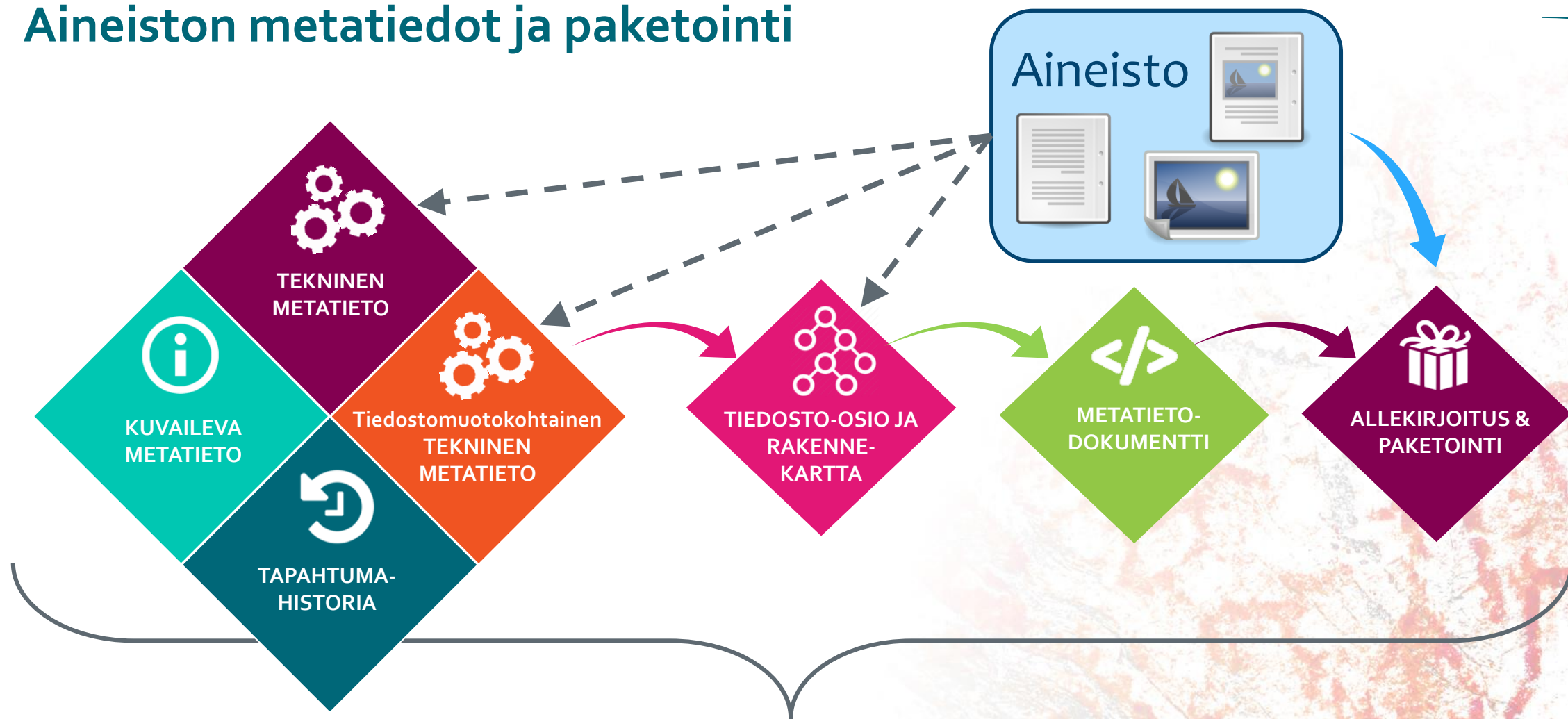
- > Sanasto
- > Hyödyntäjäksi
- > Palvelusopimukset
- > Yhteistyöryhmä
- > Liity sähköpostilistalle
- > Tietosuojaseloste
- > Saavutettavuusseloste

- > Etusivu
- > Uutiset
- > Määrittelyt
- > Julkaisut
- > Työkalut

Sähköpostiosoitte  
RSS  
Twitter  
GitHub

Copyright © 2010

# Aineiston metatiedot ja paketointi



**Paketointityökalu  
(dpres-siptools)**

Työkaluja saatavilla @ GitHub

<https://github.com/Digital-Preservation-Finland/>

# Paketointityökalu ja aineiston metatiedot



## Tekninen metatieto

Yleiset tekniset metatiedot, kuten tiedostomuodot, tarkistussummat, jne.



## Kuvaileva metatieto

Valmiin XML-muotoisen kuvailumetatiedon tuonti



## Tiedostomuotokohtainen tekninen metatieto

Kuvan, äänen, videon tai rakenteellisen tekstin ominaisuudet



## Tapahtumahistoria

Aineiston syntyhistoria ja muut aineistoon kohdistuneet tapahtumat



## Tiedosto-osio ja rakennekartta

Tiedostoluettelo, aineiston rakenne sekä metatiedon ja tiedostojen linkitykset



# Aineiston paketointi



## Metatietodokumentti

Metatiedot koostetaan yhteen dokumenttiin (määritelty koostemuoto)

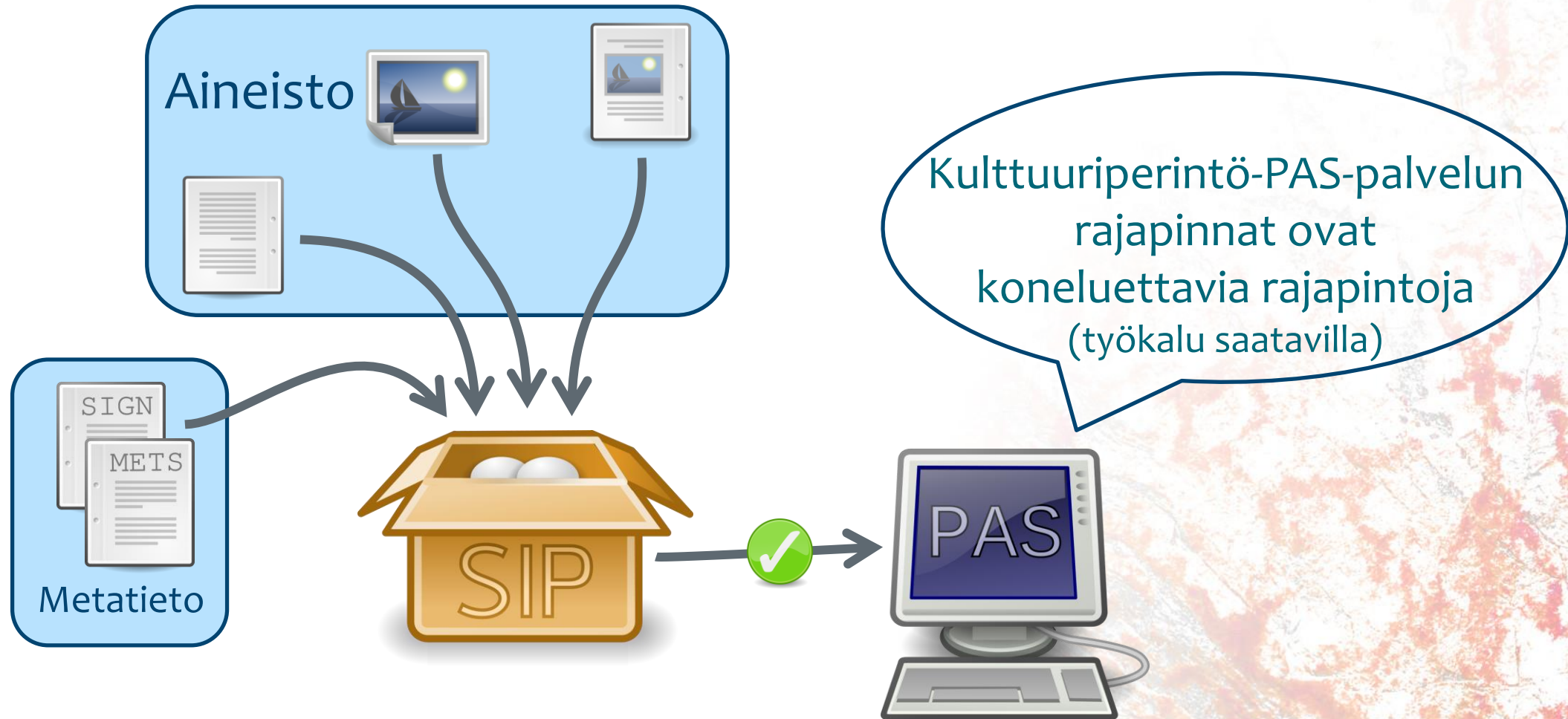


## Allekirjoitus & paketointi

Allekirjoitetaan ja paketoidaan lähetettäväksi PAS-palveluun



# Nyt paketti on valmis pitkäaikaissäilytykseen!



# Yhteistyötä organisaatioiden kanssa nyt ja tulevaisuudessa

(myös muiden kuin asiakkaiden)

- PAS-palveluiden sähköpostilista
  - <https://www.digitalpreservation.fi/2021-liity-pas-palveluiden-sahkopostilistalle>
- Kuukausittaiset #PASKaffet
  - Joka kuukauden ensimmäinen perjantai klo 14:00-14:45
- #PASKlinikat
  - Säännöllisen epäsäännöllisesti
- Koulutustilaisuuksia enemmän tai vähemmän säännöllisesti
- [twitter.com/dpres\\_FI](https://twitter.com/dpres_FI)
- Virallinen tukiosoite: [pas-support@csc.fi](mailto:pas-support@csc.fi)
- <https://www.digitalpreservation.fi>



**YOU'RE NOT  
ALONE...**

[pas-support@csc.fi](mailto:pas-support@csc.fi)