



# PAS-yhteistyöryhmä 15.6.2023



# Kokouksen avaus

- Todetaan läsnäolijat
- Esityslistan hyväksyminen
- Edellisen kokouksen muistio





# PAS-palveluiden rajapintojen uudistaminen



# PAS-palveluiden rajapinnat uudistuvat

- PAS-palveluiden rajapinnoista tulossa uusi pääversio (v3)
- Suurimmat muutokset:
  - SFTP-siirto korvataan HTTP-siirrolla
  - Siirtojen tilan seuranta ja tarkastusraporttien nouto HTTP:n kautta
  - Jakelupaketin muodostamiskutsu muuttuu
  - Uutena resurssina käsittelyhistoria
  - Uusia tilastotietoja?
- Vanhat rajapintaversiot toimivat vielä rinnakkain uuden kanssa



# Rajapintojen resurssit

- HTTP REST mallissa hallinnoidaan toimintoja resurssien kautta
- PAS-palveluiden rajapintojen resurssit:
  - **transfers** – Siirrot
  - **search** – Hakutyökalu
  - **preserved** – Säilytyspaketit
  - **disseminated** – Jakelupaketit
  - **statistics** – Tilastotiedot
  - **history** – Käsittelytiedot
  - **public-keys** – Julkiset avaimet

# Rajapintojen toiminnot

Resurssi	Metodit	Kuvaus
/transfers	POST	Aineiston siirron alustaminen
/transfers/<siirron tunniste>	HEAD, PATCH	Aineiston siirto PAS-palveluun
/transfers	GET	Siirtojen listaus
/transfers/<siirron tunniste>/status	GET	Siirron tila PAS-palvelun vastaanotossa
/transfers/<siirron tunniste>/report	GET	Siirtopaketin tarkastusraportin nouto
/search	GET	Hakutoiminto
/preserved/<aip-id>	GET	Säilytyspaketin tiedot PAS-palvelussa
/preserved/disseminate	POST	Säilytettävän aineiston muodostaminen jakelupaketiksi
/disseminated	GET	Jakelupakettien listaus
/disseminated/<dip-id>	GET	Jakelupaketin tilan PAS-palvelussa
/disseminated/<dip-id>/download	GET	Jakelupaketin nouto
/history/transfers	GET	Siirtojen käsittelytietojen listaus
/history/preserved	GET	Säilytettävän aineiston käsittelytietojen listaus
/history/disseminated	GET	Aineiston jakelun käsittelytietojen listaus
/statistics/overview	GET	Tilastotiedot
/public-key/dip	GET	Jakelutoiminnon julkisen avaimen nouto

# Rajapintojen toiminnot esiteltynä



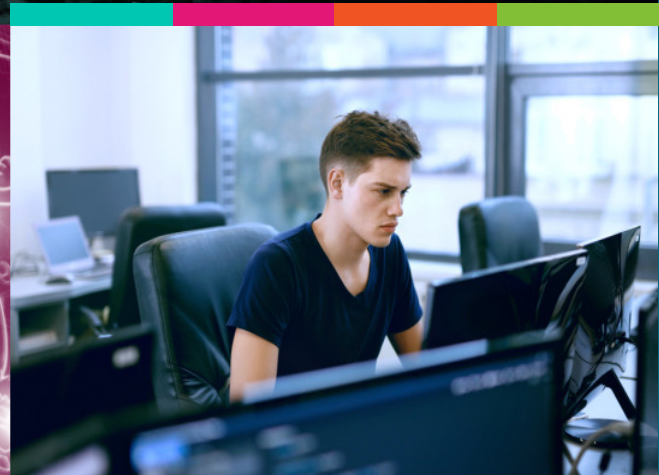


# Rajapintauudistuksen eteneminen

- Rajapinnoista luodaan uusi määrittely, jonka on tarkoitus julkaista ensi vuoden alussa
- Rajapintojen toiminnallisuus toteutetaan vaiheittain:
  - Vuoden 2024 aikana toteutetaan HTTP siirto ja siirtojen tilan seuranta
  - Vuoden 2025 aikana toteutetaan muut toiminnallisuudet
- Vanhat rajapinnat (v1 ja v2) poistetaan kun rajapintojen v3 on toteutettu
  - Suunnitelman mukaan 2026 vuoden alussa
  - Täten hyödyntävät organisaatiot voivat kerralla siirtyä v1:stä ja v2:sta v3:seen mikäli niin haluavat



# Yhteistyöryhmän syksyn 2023 toiminnan suunnittelu



# Yhteistyöryhmän syksyn 2023 toiminnan suunnittelu

Ote joulukuun 2022 kokouksen pöytäkirjasta:

*Sovittiin että painopiste vuonna 2023 on seuraavissa asioissa, normaalien vuosikellon mukaisten asioiden lisäksi:*

- *PAS-palveluiden rajapinnan uudistaminen*
- *Hyödyntäville organisaatiolle tarjottavien työkalujen kehityksessä ja niihin liittyvissä koulutuksissa*
- *Aineiston poiston PAS-palvelusta suunnittelemisen aloittaminen*
- *PAS-kokonaisarkkitehtuurityön jatkaminen*



# Yhteistyöryhmän syksyn 2023 toiminnan suunnittelu

- Vuosikellon mukaiset asiat
- PAS-kokonaisarkkitehtuuri
- Kokemukset ensimmäisestä laajasta migraatiosta (case: arc/warc-migraatio)
- PAS-palveluiden rajapintojen uudistaminen
- Aineiston poistaminen PAS-palvelusta
- AlmaLinuxiin siirtyminen
- Paketointikomponentin uudistaminen
- Työpajat / Koulutukset
- Muuta?

# PAS-palveluiden ohjelmistokerros





# PAS-palveluiden ohjelmistokerros

- Viimeinen Python2 päivitys oli maaliskuussa
- RHEL 9 / AlmaLinux 9 -siirtymä
  - On tehty ja tehdään kesän aikana
  - Asennusohjeet syksyllä
- Paketointityökalun uudistus
  - Ensivilkaisu oli PAS-seminaarissa koulutuspäivänä
  - Viedään Githubiin saataville
    - Toivottiin PAS-seminaarissa tutustumista ja palautteen antamista varten
    - Ei sisällä kaikkia ominaisuuksia
    - Ei vielä suositella tuotantokäyttöön

# PAS-palveluiden uudet ominaisuudet Q2/2023

- ARC/WARC-migraation kehitystyö
  - Erilaisia viimeistelyjä migraatioon
  - Koeaineiston kerääminen
  - Koemigraation aloitus
- Nauhamigraatioon liittyvä kehitystyö on edistynyt
- Aputyökaluja tietokantojen ylläpitämiseen
- Tiedostomuotojen analysointityökalu
  - MPEG-PS (Program Stream) version tunnistaminen
  - file- ja libreoffice-ohjelmien polkujen käsittely
  - Mediainfo-bugi virallisessa jakelussa: selvitys ja korjauspyyntö sekä väliaikainen paikkaus
- Bugikorjauksia ja muita kehitystöitä
- SAPA-kehitystyö omassa projektiryhmässään
- Fairdata PAS
  - RHEL 9 / AlmaLinux 9 –työ on edistynyt
  - Uusi paketointityökalu on edistynyt
  - Hallintaliittymä
    - Käyttöliittymäsuunnittelutyö
    - Muita pienempiä muutoksia
  - Paketointi- ja validointiprosessin muutokset
  - Säilytyspäättökohtainen autorisointiuudistus on aloitettu ja edistynyt
  - SEG-Y -tiedostomuodon tunnistus (bittitaso)
  - Bugikorjauksia ja muita kehitystöitä



# Sopimukset – KP-PAS (1/2)

(12.6.2023)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Celia	110			
Kavi	2 400			
Kansallisarkisto Kansallisarkiston vastaanottamat alkujaan digitaaliset valtiorhallinnon asiakirjalliset aineistot	41			
Kansallisarkisto VAPA-järjestelmään siirretyt tietoaineistot	1			
Kansallisarkisto Kansallisarkiston massadigitointi-hankkeen aineistot	114			
Kansallisarkisto Kansallisarkiston digitaaliarkistosta siirrettävät aineistot ja takautuvan digitoinnin aineistot	805			
Kansallisarkisto Kansallisarkiston yksinomaan digitaalisessa muodossa olevat yksityisarkistoaineistot	27			

# Sopimukset – KP-PAS (2/2)

(12.6.2023)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Kansallisgalleria	20			
Kansalliskirjasto Kulttuuriaineistolain nojalla kerätyt aineistot	355			
Kansalliskirjasto Kansalliskirjaston digitoimat kulttuuriperintöaineistot	175			
KOTUS	60			
Museovirasto	1			
Musiikkiarkisto	70			
Svenska litteratursällskapet i Finland	50			
Yhteiskuntatieteellinen tietoarkisto	1			
<b>Yhteensä</b>	<b>4 230,00</b>			



# Säilytyspäätökset – FD-PAS (1/2)

(12.6.2023)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Geologian Tutkimuskeskus GTK:n tomografialaitteen tuottamat tietoaaineistot	12			
Helsingin yliopisto Helsingin yliopiston SMEAR-aineistojen valikoima meteorologisia - ja ilmanlaatumittauksia	2			
Helsingin yliopisto M. cinxia and C. melitaeorum in the Åland metapopulation system	2			
Helsingin yliopisto FIRE (The Finnish Reflection Experiment)	1			
Helsingin yliopisto Luomuksen aineistot	150			
Itä-Suomen yliopisto SENSOTRA	1			

# Säilytyspäätökset – FD-PAS (2/2)

(12.6.2023)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Jyväskylän yliopiston kiihdytinlaboratorio 250-Nobeliumin hajoamisspektroskopia	1			
Oulun yliopisto, Sodankylän geofysikaalinen observatorio <u>Havaintoaineistot</u>	30			
Tampereen yliopisto Yhteiskuntatieteiden tiedekunnan Kansanperinteen arkiston A-K-kokoelma	2			
Turun yliopisto Historian, kulttuurin ja taiteiden tutkimuksen arkiston aineistot (HKT-arkisto)	20			
Åbo Akademi Samlingar vid Åbo Akademis bibliotek	10			
<b>Yhteensä</b>	<b>231</b>			



# Ilmoitusasiat

- PAS-määrittelyiden päivittäminen – kysely käynnissä 30.8.2023 asti
  - <https://link.webpolsurveys.com/S/212D3C71B2F6BB29>
- PAS-seminaari järjestettiin 24.-25.4.2023 CSC:llä
- CSC järjesti OPF:n vuosikokouksen (AGM) ja teknisen ohjausryhmän (OAG) kokoukset 24.-25.5.2023 Oodissa
- iPPRES 2023 -konferenssiin hyväksytty
  - *Calculating the Carbon Footprint of Digital Preservation - A Case Study*
- Aalto-yliopiston Tutkimuspalvelupäivät 21.-23.8.2023
  - *Calculating the Carbon Footprint of Digital Preservation*
  - *On Preserving Research Data – What to Consider Before Ingesting Data*

# Kokouksen päättäminen

- Muut asiat
- Seuraava kokous
- Kokouksen päättäminen

