



PAS-yhteistyöryhmä 17.6.2021



Kokouksen avaus

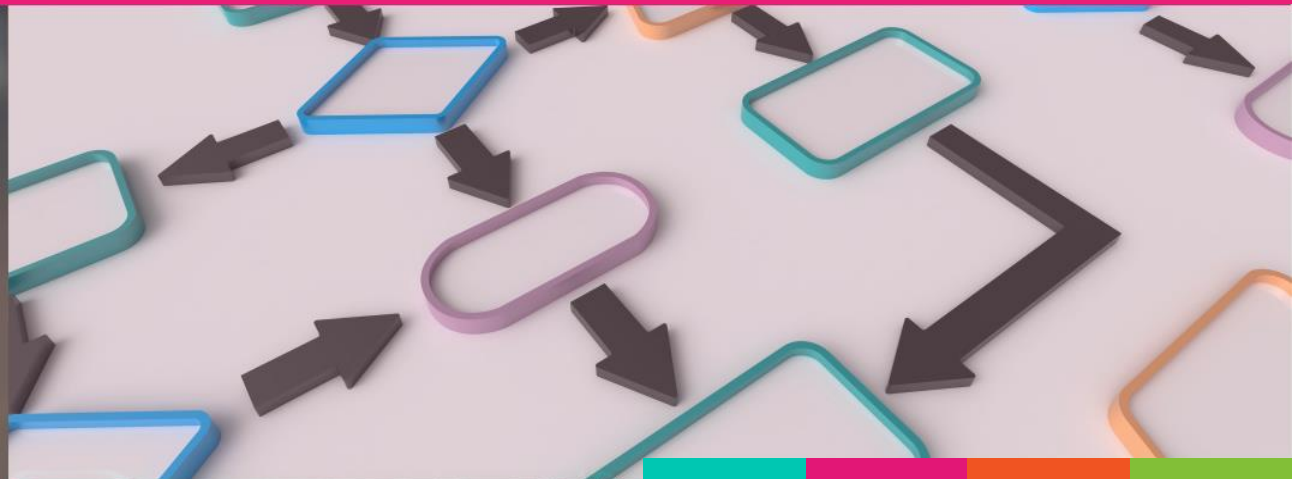
- Todetaan läsnäolijat
- Esityslistan hyväksyminen
- Edellisen kokouksen muistio



Sähköinen arkistointi ja todistusvoimaisuus (Markus Merenmies)



METS2 suunnitelmien kuulumisia



Alustava katsaus METS2:een

- METS Editorial Board suunnittelee uutta versiota METS XML -määrittelystä
- Ajatuksena on tehdä METS:stä helpommin lähestyttävä, joustavampi, korjata jotkin sen tekniset ongelmakohdat sekä siivota tarpeeton pois
 - Boardissa on alun perin keskusteltu sekä ns. xlink-skeeman ongelmista että helpotetun METS:in tarpeesta ("METS Light"), joihin METS2 pyrkii vastaamaan
- Uusi määritelmä ei ole vanhan kanssa yhteensopiva
- Ei vielä julkistettu kommenteille
- Kyseessä on pitkälti tekninen muutos, joka kuitenkin mahdollistaa joustavamman METS:n käytön
 - Päätasot ja erilaiset yksittäiset asiat muuttuvat
- Päätasojen sisällä ollaan silti paljonkin nykyisen METS:n mukaisessa maailmassa

METS2: Keskeisiä suunniteltuja muutoksia

- Päätasoja voi olla kutakin korkeintaan yksi:
 - Otsake-osio
 - Metatietojen osio
 - Tiedostojen osio
 - Rakennekartta-osio (rakennekartta muuttunee vapaaehtoiseksi)
 - (nykyisen määrittelyn `<structLink>` ja `<behaviorSec>` poistetaan)
- Metatiedon laji (kuvaileva, tekninen, provenanssi jne.) vapautetaan
 - Mahdollistaa metatietolohkojen joustavamman käytön, esim.
 - Hybridimetatietomuotojen ilmaisu
 - Yksityiskohtaisempi metatietojaottelu (esim. säilytysmetatieto erikseen)
 - Nykyisessä METS:ssä näille kullekin lajille on eri elementtinsä
 - METS2:ssa on yleinen (toistettava) metatietolohko, jonka laji voidaan määritellä attribuutilla.
 - Metatietolohkoja voisi myös ryhmitellä
 - Suositussanasto

METS2: Keskeisiä suunniteltuja muutoksia

- Kontrolloidut sanastot vapautetaan
 - Nykyisessä METS:ssä on paljon attribuutteja, joissa on kontrolloitu sanasto, ja sitten näille on toiset erilliset samaa asiaa tarkoittavat attribuutit muille termeille
 - Suositussanastot luodaan
 - Hyödyntäjän ei tarvitse enää muistaa, onko termi virallisessa sanastossa vai ei
 - Ohjelmakoodi tietojen lukemiseksi ja kirjoittamiseksi voi siistiytyä
- Tietyt sisäkkäiset elementit poistetaan
 - esimerkiksi tiedostoryhmien alle ei voisi muodostaa enää alaryhmiä, jolloin rakenteellinen tieto ohjataan rakennekarttaan
- Tiedosto-osiossa on vielä vähän auki jotkin yksityiskohdat
- XLINK-laajennoksesta luopuminen

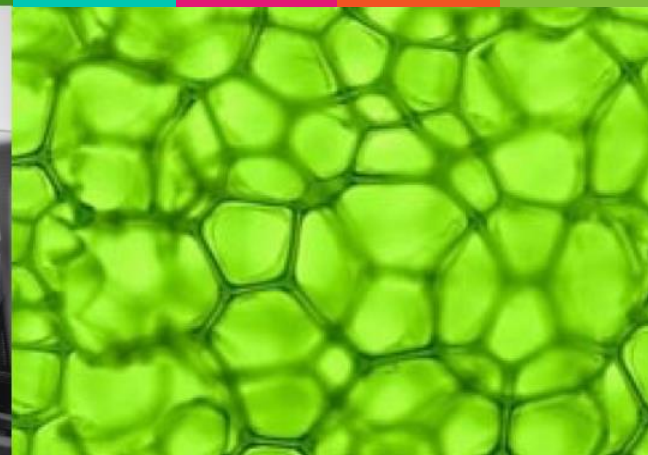
METS2: Huomioita

- Uudessa standardissa elementtien tasot menevät vieläkin aika syvälle, eli elementtejä on sisäkkäin vieläkin aika paljon, mikä tuntuu hankalalta
- Binääridatan upottaminen METS:ssä säilyy, mutta sen tarve voi olla kyseenalainen. Jos sen poistaisi, niin METS voisi vähän yksinkertaistua.
- Suomen kansallisia laajennoksia ei tulla sellaisenaan avaamaan yleiseen käyttöön
 - Sen sijaan näissä on havaittu eri käytäntöjä eri ympyröissä, jolloin on nähty tarve yhteiselle suositukselle, mutta mieluummin jo olemassa olevilla elementeillä ja attribuuteilla
 - Kasvavan aineiston linkitykset on tunnistettu ainakin osittain ongelmaksi
 - Paikallisten laajennoksien mahdollisuus säilytetään
- Maailmassa on olemassa nykyiseen METS:iin pohjautuvia muita kv-standardeja

METS2 ja PAS

- Kun määrittely tulee yleisesti kommentoitavaksi, niin siihen kannattaa antaa kommentteja
 - Kun määrittely ei edes yritä olla taaksepäin yhteensopiva, niin se on tilaisuus tehdä muutoksia reippaammin, toisaalta reippaampia muutoksia kannattaa arvioida huolella
- Kannattaa seurata, kuinka METS2 otetaan maailmalla vastaan
- METS1 tulee olemaan käytössä vielä pitkään
- Toteutus PAS:iin olisi laaja ja koskettaisi kaikkia toimijoita muutostyön muodossa. Kun nyt on olemassa paketoitumäärittely 1.7.3, niin METS2 käyttöönotto muuttaisi versionumeroksi selkeästi 2.0.0
 - Muutostarpeelle pitää olla selkeät perusteet
 - Ohjelmistokehitystyö olisi laaja, paketoitukomponentti tosin auttaisi hyödyntäviä organisaatioita
 - Nykyinen PAS-paketoitumäärittely on määrittelytyönä sovitettavissa METS2:een aikalailla suoraviivaisesti, joskin erilliset suositukset joillekin asioille käsiteltävä erikseen
 - Joustavammat ominaisuudet mahdollistavat paketoitumäärittelyn uudistamisen

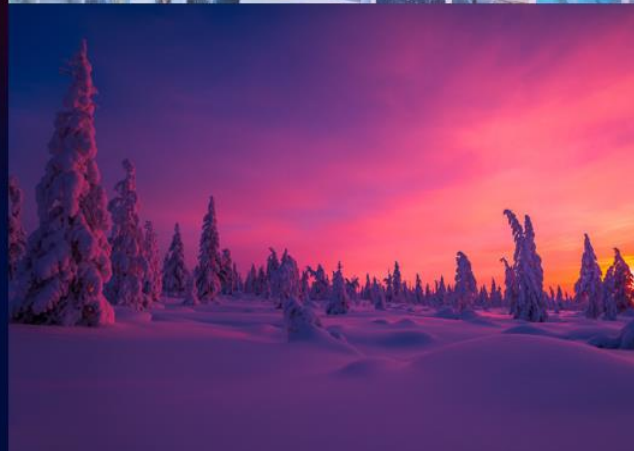
Yhteistyöryhmän syksyn teemat



Yhteistyöryhmän syksyn teemat

- Vuosikellon mukaiset asiat
- PAS-palvelun rajapinnat työpaja
- Kokemukset ensimmäisestä laajasta migraatiosta (case: arc/warc-migraatio)
- Loogisen säilyttämisen tiekartta
- Valittujen kansainvälisten määrittelyjen suhde kansallisiin PAS-palveluihin

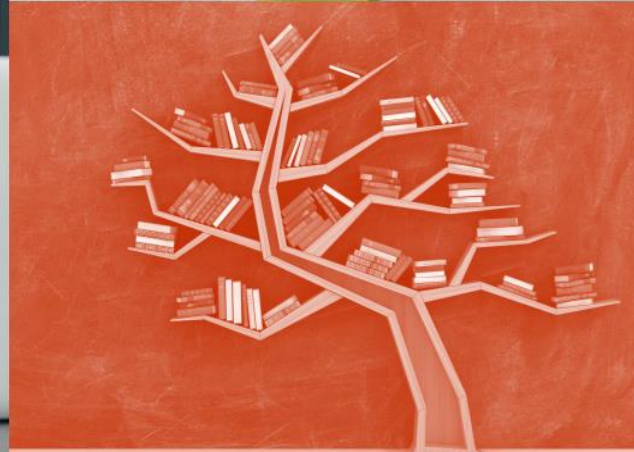
PAS-palveluiden uudet ominaisuudet Q2/2021



PAS-palveluiden uudet ominaisuudet Q2/2021

- Säännöllinen eheydentarkastus maantieteellisesti hajautetussa ympäristössä
- REST-rajapinnassa usean käyttäjätunnuksen tuki samalle sopimustunnisteelle
- Validointiraportin muodostumisen nopeuttaminen
- ARC-tiedostomuototuen poisto vastaanotosta
- PRONOM-tiedostomuotokoodien ja tiedostomuotoversioiden päivitys
- Virustarkastuksen korjaus suurille tiedostoille
- Muita pienempiä bugikorjauksia
- SAPA-kehitystyö omassa projektiryhmässään erillisen suunnitelman mukaisesti
- Fairdata PAS:
 - Siirtyminen Python3-ympäristöön
 - Mahdollisuus demoympäristöön
 - Integraatiopäivitys muihin Fairdata-palveluihin
 - Paketointipalvelun tuki kaikille PAS-palveluun toteutetuille tiedostomuodoille

PAS-palveluiden säilytyksen tilanne ja käynnissä olevat käyttöönotot



Sopimukset – KP-PAS (1/2)

(9.6.2021)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Celia	110	64,81	58,92 %	23 559
Kavi	1200	1 002,52	83,54 %	2 672
Kansallisarkisto Kansallisarkiston vastaanottamat alkujaan digitaaliset valtionhallinnon asiakirjalliset aineistot	41	0,28	0,68 %	408
Kansallisarkisto VAPA-järjestelmään siirretyt tietoaaineistot	1	0,14	14,00 %	389
Kansallisarkisto Kansallisarkiston massadigitointi-hankkeen aineistot	114	31,65	27,76 %	26 834
Kansallisarkisto Kansallisarkiston digitaaliarkistosta siirrettävät aineistot ja takautuvan digitoinnin aineistot	805	90,30	11,22 %	156 343
Kansallisarkisto Kansallisarkiston yksinomaan digitaalisessa muodossa olevat yksityisarkistoaaineistot	27	0	0,00 %	0

Sopimukset – KP-PAS (2/2)

(9.6.2021)



Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Kansallisgalleria	20	0,76	3,80 %	238
Kansalliskirjasto Kulttuuriaineistolain nojalla kerätyt aineistot	165	145,01	87,88 %	1 374 935
Kansalliskirjasto Kansalliskirjaston digitoimat kulttuuriperintöaineistot	175	14,73	8,42 %	3 889
KOTUS	60	9,71	16,18 %	403
Museovirasto	1	0,23	23,00 %	20 719
Musiikkiarkisto	70	0	0,00 %	0
Svenska litteratursällskapet i Finland	50	1,14	2,28 %	166
Yhteiskuntatieteellinen tietoarkisto	1	0,01	1,00 %	1 604
Yhteensä	2 840,00	1 361,29	47,93 %	1 612 164

Säilytyspäätökset – FD-PAS

(9.6.2021)



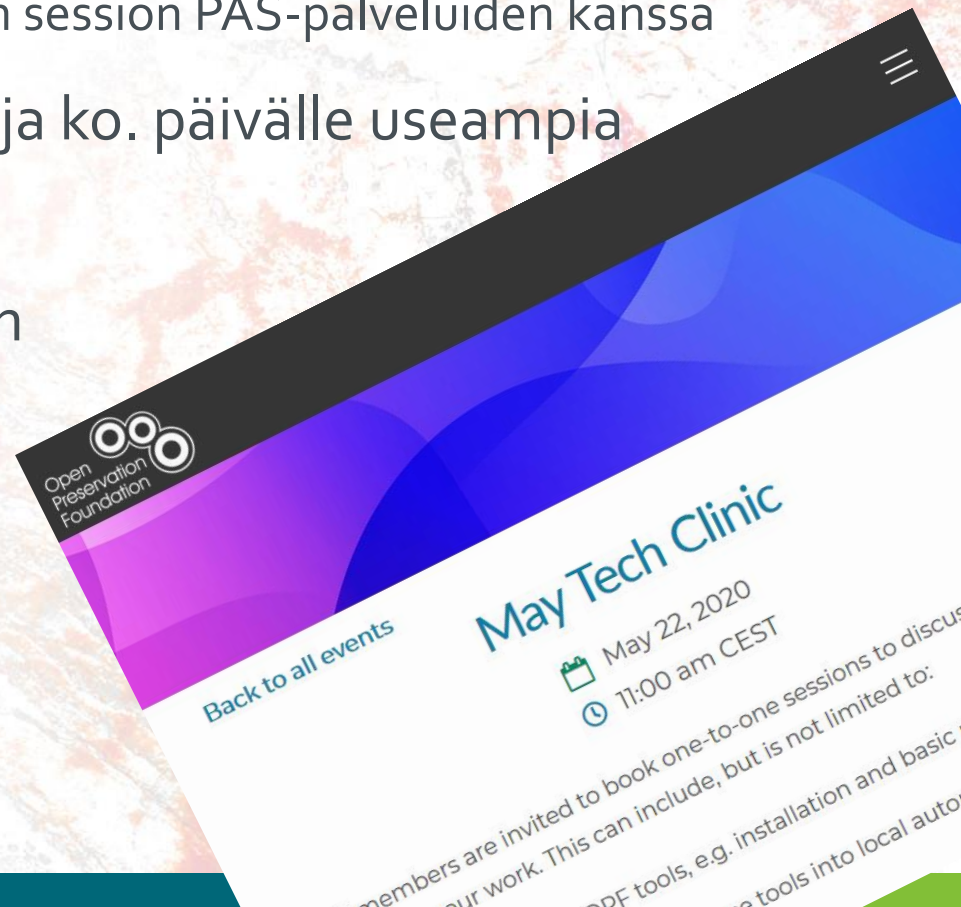
Organisaatio	Kapasiteetti (Tt)	Aineistoa (Tt)	Täyttöaste	Säilytyspaketteja
Geologian Tutkimuskeskus GTK:n tomografialaitteen tuottamat tietoaineistot	12	2,72	22,67 %	68
Helsingin yliopisto Helsingin yliopiston SMEAR-aineistojen valikoima meteorologisia - ja ilmanlaatumittauksia	2	0,01	0,55 %	13
Helsingin yliopisto M. cinxia and C. melitaeorum in the Åland metapopulation system	2	0,00	0,05 %	1
Jyväskylän yliopiston kiihdytinlaboratorio 250-Nobeliumin hajoamisspektroskopia	1	0	0 %	0
Itä-Suomen yliopisto SENSOTRA	1	0	0 %	0
Oulun yliopisto, Sodankylän geofysikaalinen observatorio Havaintoaineistot	30	0	0 %	0
Turun yliopisto Historian, kulttuurin ja taiteiden tutkimuksen arkiston aineistot (HKT-arkisto)	20	0	0 %	0
Yhteensä	68	2,73	4,01 %	82

Ilmoitusasiat

- ISO 27001 -sertifikaatti
- Määrittelyjen päivittäminen kysely käynnissä
- Tapahtumat koulutus järjestettiin 22.4.2021. Palaute positiivista
- OPF webinaari: [The Danish National Archives presents a new Concept Model for Development of Preservation Plans](#), 29.6.2021
- #PASKlinikka

#PASKlinikka

- Tarkoitus kokeilla konseptia, joka on käytössä OPF:ssä
- Matalan kynnyksen keskustelutilaisuus organisaation ja PAS-palveluiden kanssa
 - Organisaatio voi varata tietylle päivälle yksityisen esim. 45min session PAS-palveluiden kanssa
- Päivä jolloin sessioita järjestetään sovitaan etukäteen ja ko. päivälle useampia sessioita. Päiviä esim. yksi / kvartaali.
- Ei tarkoitettu "isoille" asioille; niiden suhteen jatketaan kuten tähänkin asti
- Ei korvaa #PASKaffeja, mutta täydentänee niitä
- Ensimmäiset #PASKlinikat syyskuussa



Kokouksen päättäminen

- Muut asiat
- Seuraavat kokoukset
 - Sovittu aiemmin doodlella: 21.9.2021 & 14.12.2021
- Kokouksen päättäminen

