



PAS-yhteistyöryhmä

12.9.2018



CSC – Suomalainen tutkimuksen, koulutuksen, kulttuurin ja julkishallinnon ICT-osaamiskeskus

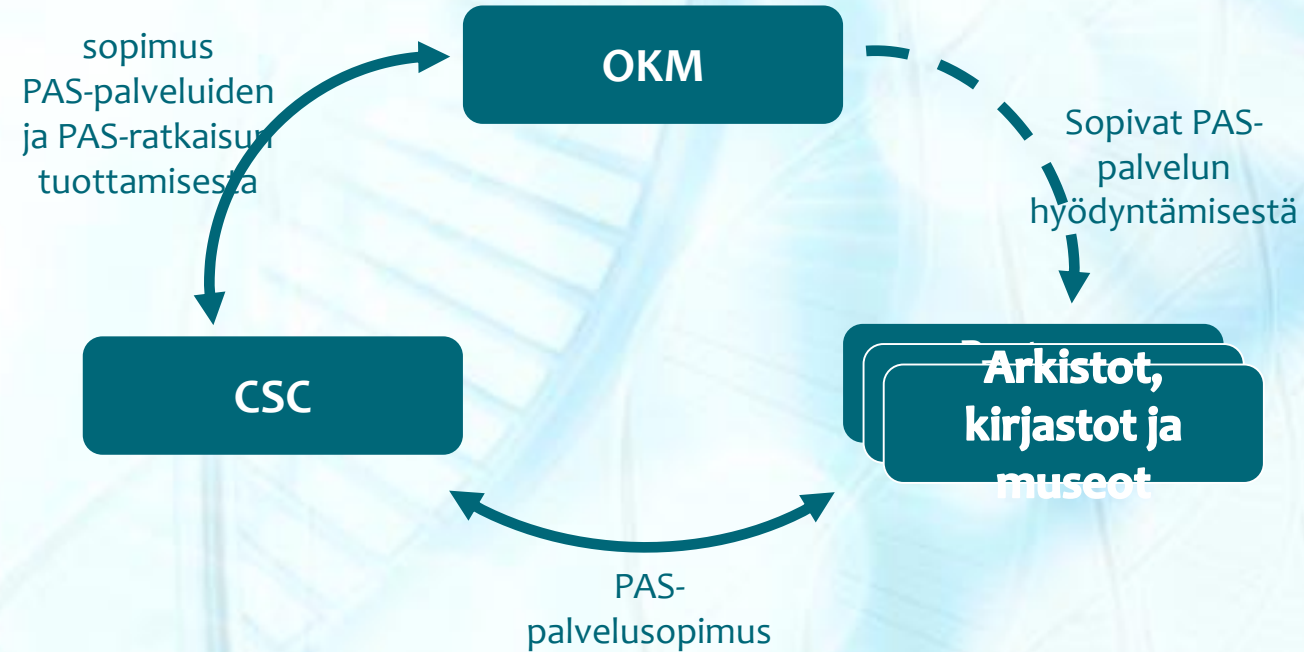
Kokouksen avaus

- Todetaan läsnäolijat
- Esityslistan hyväksyminen

Welcome

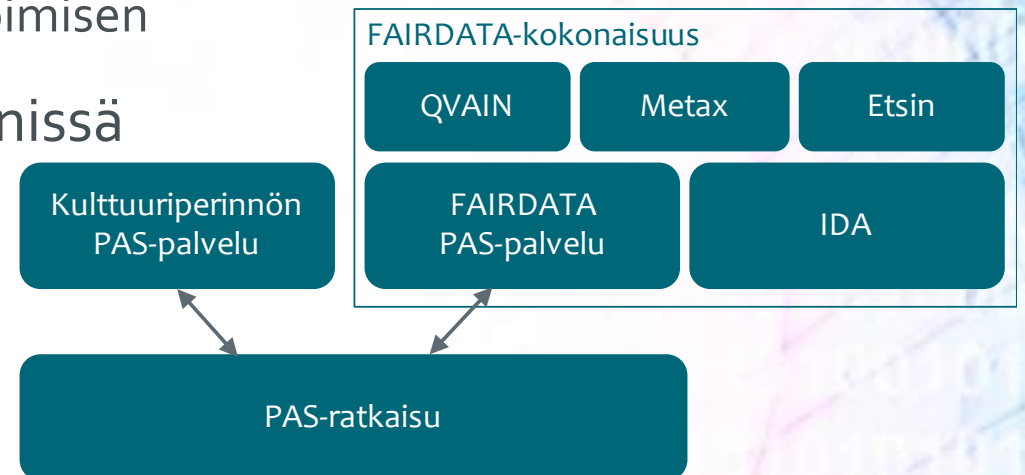
Katsaus PAS-sopimusmallin uudistamiseen: kertausta

- PAS-palvelut ja PAS-ratkaisu tuotetaan OKM:n ja CSC:n välisen "PAS-sopimuksen" pohjalta
- OKM ja hyödyntävä organisaatio sopivat PAS-palvelun hyödyntämisestä
- => hyödyntävä organisaatio ja CSC laativat PAS-palvelusopimuksen
- Kokonaisuudessa uudistustarpeita mm. KDK- ja ATT-hankkeiden päättymisen, GDPR:n ja tutkimusaineistojen säilyttämisen (Fairdata-PAS) aloittamisen vuoksi
- Toiminnan perusluonne pidetään samana
 - Palvelun maksuttomuus hyödyntäville organisaatioille säilyy jne.



Uudistuksen tilanne

- PAS-palveluiden tietosuojaseloste julkaistu
 - Yhteinen PAS-palveluille, PAS-ratkaisulle ja Fairdata-kokonaisuudelle
 - OKM rekisterinpitäjä, CSC käsittelijä
 - <http://digitalpreservation.fi/tietosuojaseloste>
- Kulttuuriperintö-PAS-palvelusopimus (luonnos) liitteineen on lähetetty kaikille hyödyntäville organisaatiolle 23.-25.5.2018
 - Sisältää KDK-hankkeen päättymisen ja GDPR:n huomioimisen
- PAS-sopimuksen viimeistely OKM:n kanssa käynnissä
- (Fairdata-PAS-palvelusopimuksen rakentaminen käynnissä)



Viimeistelyssä tarkistettavat asiat

- Miten huomioidaan aineistoihin mahdollisesti sisältyvät henkilötiedot?
 - Luonnoksessa ehdotetaan näissä tapauksissa erillisen "Käsittelytoimien kuvaus"-liitteen täyttämistä
- Miten varmistetaan KP- ja FD-palvelusopimuksien riittävä yhteismitallisuus?
 - Ehto kustannustehokkaalle palvelutuotannolle

PAS-määrittysten päivittäminen

- Webropol kysely (5.7-31.8.2018)
 - 6 vastausta (tuplaantui viime vuodesta 😊)
 - Palautetta saatu myös käyttöönotoissa ja FAIRDATA-PAS palvelun kehityksestä
- Dokumenttien päivitys (CSC)
 - Päivitetyt versiot yhteistyörhmälle kommenteille 30.11.2018 mennessä
- Hyväksyminen
 - Yhteistyörhmä (12.11.2018)
- Julkaisu
 - Alkuvuodesta, heti kun muutokset on toteutettu tuotantoon ja skeemoihin

PAS-määrittysten päivittäminen

Yleistä

- Tuki paketoitumäärityksen versiolle 1.4. voitaneen lopettaa?
 - Versiota ei kukaan enää käytä ja uusille organisaatioille vain 1.7 on mahdollinen
- Määrityksen METS xml:lang attribuuteista voitaneen poistaa vaatimus ISO 639-2:lle ja korvata se xml:lang attribuutin virallisella määrityksellä (BCP 47)
- Määrityksiin lisätään tarkennus text/plain tiedostomuodon käytöstä
- Esitetty toive kansallisen METS:n ja EARK-METS:n yhteensopivuuden selvittämiseksi
 - Yhteensopivuuden saavuttaminen näyttää mahdottomalta
- Osa kommentteista liittyi ohjeistukseen, eikä siten kuulu määrityksiin
 - Tarpeen vaatiessa kysymme lisätietoja kommenttien antajilta

PAS-määrittelysten päivittäminen

Metatiedot

- METS: Lisätään tuki versiolle 1.12
 - Yhteensopiva version 1.11 kanssa; agent/note:een lisätty anyAttribute##other
 - Ei vaikutuksia määrittelyyn eikä toteutukseen hyödyntävissä organisaatioissa.
- PREMIS: Tuki versiolle 3.0?
 - 3.0 ei ole yhteensopiva 2.X kanssa. Samanaikainen tuki haasteellista.
 - Muutos 2.X => 3.0 vaikuttaa myös esim. validointiraportteihin ja siten tuotannossa oleviin hyödyntäviin organisaatioihin.
 - Onko PREMISin versiolle 3.0 konkreettista tarvetta?

PAS-määrittysten päivittäminen

Kuvailevat metatiedot

- DataCite
 - Versio 4.1 lisätään määrittelykseen. Käytännötilanne jo nyt tuotannossa.
 - Tuki 4.0:lle poistetaan, koska versio ei ole yhteensopiva version 4.1 kanssa.
- EAD₃
 - Versio 1.1.0 lisätään määrittelykseen
- MODS
 - Versio 3.7 lisätään määrittelykseen
- VRA Core 4.0
 - Täsmennetään määrittelykseen, että tuetaan "restricted" versiota (käytäntö jo nyt)

PAS-määrittysten päivittäminen

Tiedostomuodot

- Siard 2.0 lisätään säilytyskelppisiin
 - Tarve sekä kulttuuri- & tutkimus PAS-palveluissa
- MatLab & HDF5 lisätään säilytyskelppisiin (rajoituksilla)
- SVG ja FFV1 lisätään todennäköisesti säilytyskelppisiin
 - Evaluointi tekemättä
- Apple Proresin osalta selvitetään mahdollinen migraatio johonkin toiseen tiedostomuotoon
 - Suljettu tiedostomuoto
- DV & MPEG-2 ehdotettu siirrettäväksi siirtokelppisistä säilytyskelppisiin
 - Tiedostomuodot määrittelyn tarkoituksena on myös ohjata mitä tiedostomuotoja kannattaa käyttää ja tämä muutos saattaisi ohjata organisaatioita väärään suuntaan

PAS-määrittysten päivittäminen

Rajapinnat – Tilastot



PAS-määrittysten päivittäminen

Rajapinnat - Tilastot

- Muita ehdotuksia tilastoitaviksi suureiksi
 - Migraatioiden määrä
 - Migroidujen säilytyspakettien määrä
 - Poistettujen säilytyspakettien määrä
 - Säilytyksessä oleva kapasiteetti per tiedostomuoto
- Toiminnallisuus tilastointiin liittyen
 - Kertymien/tilastojen tarkasteleminen aikajaksoittain

PAS-määrittysten päivittäminen

Rajapinnat - Toiminnallisuus

- Aineistojen poistaminen
 - Tulossa, mutta ei korkealla prioriteetilla
- Mahdollisuus vastaanoton työkulun parempaan seurantaan
 - Tulossa; työnalla
- Säilytyspaketin toimittaminen suoraan hyödyntävälle organisaatiolle
 - Rikkoo OAIS-mallin
 - Haasteellinen myös siksi, että säilytyspaketti ei ole yksittäinen kokonaisuus. Esim. säilytysmetatiedot ovat erillisessä tietokannassa joita ei fyysiseen säilytyspakettiin päivitetä niiden muuttuessa

PAS-määrittysten päivittäminen

Natiivitiedostojen käsittelyn tarkentaminen

- Erityisesti FAIRDATA-PASsissa on tarvetta säilyttää tiedostomuotoja, jotka eivät ole standardoituja tai mutta mahdollisesti paljon käytettyjä tietyllä tutkimusalalla
- PAS-palveluun ei mitä tahansa tiedostomuotoja voida hyväksyä, joten pitää kehittää yhteiset pelisäännöt,
 - Mitä näissä tapauksissa tehdään?
 - Millä ehdoilla?
- Joka tapauksessa, tällaisille tiedostomuodoille voidaan
 - taata ainoastaan bittitason säilytys
 - ja asiasta on aina sovittava tapauskohtaisesti

PAS-määrittysten päivittäminen

Natiivitiedostojen käsittelyn tarkentaminen

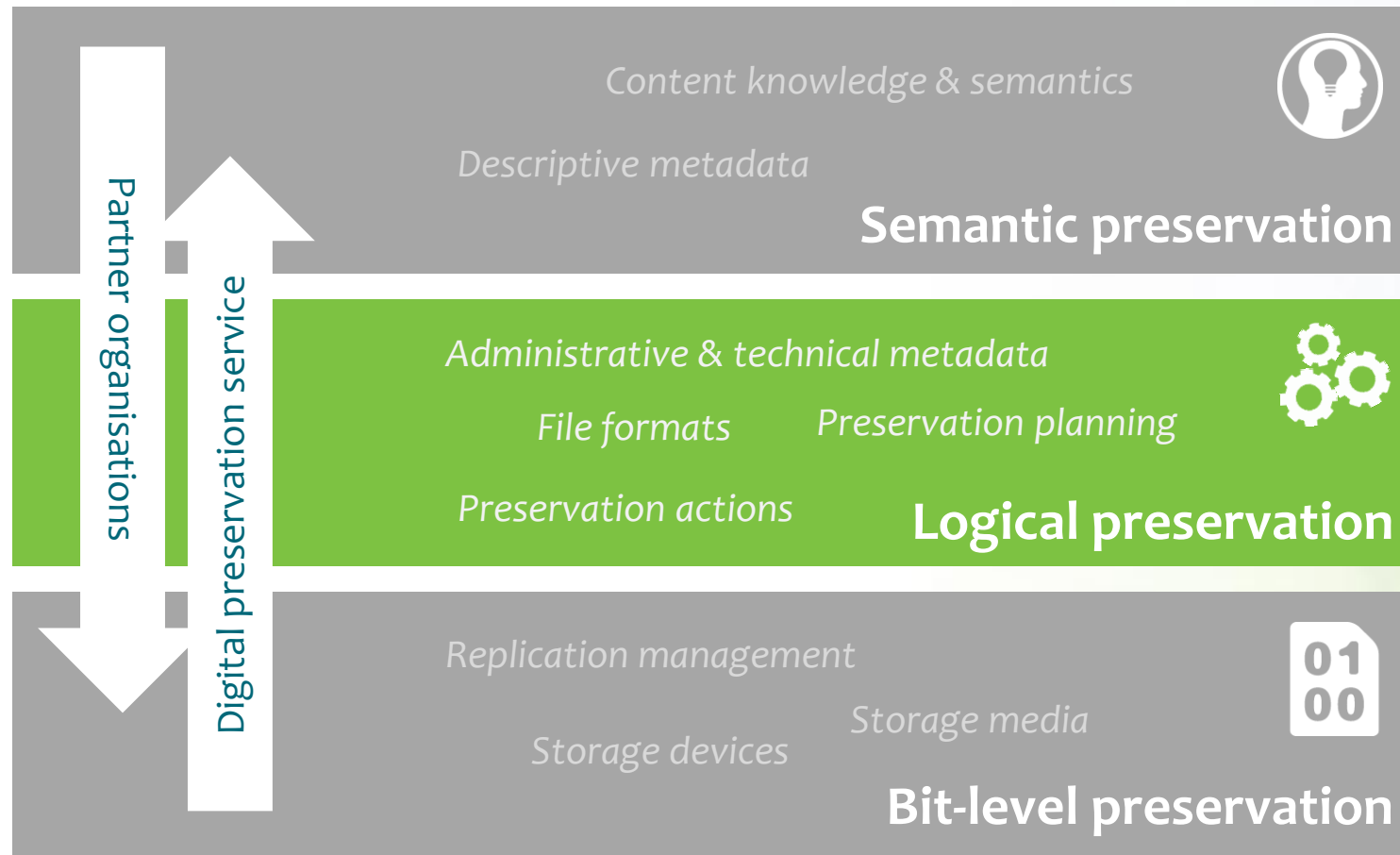
- Vaaditaan...
 - Tiedostomuodon yhteyteen lisättävä säilytyskelvoinen tiedostomuoto?
 - jos se on kohtuullisella vaivalla tehtävissä informaatiota kohtuuttomasti hävittämättä?
 - Tiedostomuoto pitää pystyä identifioimaan koneellisesti?
 - PRONOM tunniste (tai vähintään hakemus tehty)?
 - Tiedostomuodon metadata vaatimukset?
 - Eli mitä ja miten tiedostomuodosta pitäisi kertoa jos tiedostomuodon määrittäjiä ei ole olemassa?
 - Dokumentaatio (jos sellainen on saatavilla)?
 - Validaattori?

PAS-määrittysten päivittäminen

Pysyvät tunnisteet määriyksille

- Pysyvät linkit määrittysten uusimpiin versioihin
 - W3C-malli, eli cool URI, joka on joko varsinainen osoite tai uudelleenohjataan siihen
 - pysyvä tunnus (DOI tai URN), joka resolvoidaan esimerkiksi speksin uusimman version osoitteeksi tai osoitteeksi, joka uudelleenohjataan varsinaiseen sijaintipaikkaan
- Määrittysten kaikkien versioiden arkistointi

Loogisen säilyttämisen toiminnallisten vaatimusten täsmentäminen



Loogisen säilyttämisen toiminnallisten vaatimusten täsmentäminen

- Jakaudutaan 2-3 ryhmään pohtimaan loogisen säilyttämisen toiminnallisia vaatimuksia
- Ryhmätyön tarkoituksena on
 - Priorisoida annetut vaatimukset kolmeen luokkaan toivotun toteutus järjestyksen mukaisesti
 - Pyrkikää jakamaan vaatimukset luokkiin mahdollisimman tasaisesti
 - Lisätä kunkin vaatimuksen alle muutamia ajatuksia mitä ko. toiminnallisuus ryhmän mielestä pitää sisällään
 - Keksiä lisää toiminnallisia vaatimuksia

Loogisen säilyttämisen toiminnallisten vaatimusten täsmentäminen

- Aineiston piirteiden tunnistaminen ja priorisointi säilytyksen suunnittelussa
- Digitointiohjeiden ylläpitäminen (tiedostomuotomäärityksen laajuudessa)
- Migraatioiden hyväksyntä -työkalu hyödyntäville organisaatioille
- Migraatioiden laadunvarmistus
- Migraatioiden suunnittelu -työkalu hyödyntäville organisaatioille
- Migraatioiden toteuttaminen siirtokelpoisille tiedostomuodoille
- Migraatioiden toteuttamisen (julkinen) raportointi ja seuranta
- Migraatiopolkujen määrittäminen siirtokelpoisille tiedostomuodoille
- Natiivitiedostojen käsittely ja vaatimusmäärittely
- Siirtokelpoisten tiedostomuotojen määrityksen ylläpitäminen

Loogisen säilyttämisen toiminnallisten vaatimusten täsmentäminen

- Säilytyskelpoisten tiedostomuotojen määrityksen ylläpitäminen
- Säilytyssuunnitelman loogisen tason määrittäminen
- Säilytyssuunnitelman teknisen tason toteuttaminen
- Teknisten metatietojen määrityksen ylläpitäminen
- Tiedostomuotojen validointi
- Tiedostomuotokohtainen raportointi/näkymä
- Tiedostomuotokohtaisten ominaisuuksien kartoittaminen
- Tiedostomuotokohtaisten ominaisuuksien vaikutukset migraatioihin
- Toiminnallisuus migraatioiden tunnistamiseksi ja ennakoimiseksi
- Virheellisten tiedostomuotojen käsittely

Työpajat (ehdotukset)

- Migraatiot
- Tiedostomuotojen laadun varmistaminen
- Elävän kuvan tiedostomuodot
- Hands-on työpajoja
 - Esim. paketoitukomponentti

Käynnissä olevat käyttöönotot

- Museovirasto: Testaus testiympäristössä käynnissä
- Kansalliskirjasto (2. sopimus): Testaus testiympäristössä käynnissä
- Kotus: Testaus testiympäristössä käynnissä
- SLS: 1.vaiheen kartoitus on tehty

PAS-palvelun säilytyksen tilanne (5.9.2018)

- Säilytyksessä n.134teratavua
- Aineistopaketteja 1 200 234

Kokouksen päättäminen

- Muut asiat
- Seuraava kokous
 - 12.12.2018 klo 12:30
- Kokouksen päättäminen