

28.3.2019

Tietoarkkitehtuuriryhmän kokous

Aika Torstai 28.3.2019, klo 10.15–11.45

Paikka Kansalliskirjasto, Fabianian auditorio (C221), Yliopistonkatu 1, Helsinki

Jäsenet

Kehittämispäällikkö Kenneth Ahlfors, Kansallisarkisto
Ylitarkastaja Miia Herrala, Kansallisarkisto (puheenjohtaja)
Ylitarkastaja Minna-Liisa Mäkiranta, Kansallisarkisto - etäyhteydellä
Erikoissuunnittelija Tuomo Häkkinen, Kansallinen audiovisuaalinen instituutti
Koordinaattori Sampsa Heinonen, Suomen museoliitto
Senior Technology Coordinator Heikki Helin, CSC
Erikoisasiantuntija Ulriikka Puura, Kotus
Järjestelmäkehityspäällikkö Niklas Liljestrand, Svenska Litteratursällskapet i Finland
Tietoarkkitehti Katja Moilanen, Yhteiskuntatieteellinen tietoarkisto
Päällikkö Heljä Kajo, Yle Arkisto
Tietojärjestelmäpäällikkö Matias Frosterus, Kansalliskirjasto
Erikoisasiantuntija Juha Hakala, Kansalliskirjasto
Tietojärjestelmäpäällikkö Nina Hyvönen, Kansalliskirjasto
Kehittämispäällikkö Mikko Lappalainen, Kansalliskirjasto
Kehittämispäällikkö Marja-Liisa Seppälä, Kansalliskirjasto
Suunnittelija Piia Naukkarinen, Kansalliskirjasto

Suunnittelija Jaana Routakangas, Kansalliskirjasto (sihteeri)

Poissa Kirjastopalvelujen johtaja Katri Vääntinen, Helsingin kaupunki
Tietojärjestelmäasiantuntija Jari Vihtari, Kotus
Intendentti Hannu Häkkinen, Museovirasto

1. Kokouksen avaus, osallistujien toteaminen ja asialistan hyväksyminen

- Lauri Saarikoski (Yle Arkisto) on työvapaalla kevään ajan. Häntä sijaistaa tietoarkkitehtuuriryhmässä Heljä Kajo (Yle Arkisto).

Puheenjohtaja avasi kokouksen klo 10.17 ja totesi osallistujat. Asialista hyväksyttiin lisäyksellä muihin asioihin: Kirjastoverkkopäivien suunnittelu aloitettu.

2. Pitkäaikaissäilytys (Juha Hakala, Heikki Helin)

- Tarkennukset digime-ta-ryhmän työsuunnitelmaan, muutosten läpikäynti ja hyväksyntä
- PAS-yhteistyöryhmän toiminnan esittely (Heikki Helin)
- PAS-metatieto (Juha Hakala)

Tietoarkkitehtuuriryhmän työsuunnitelmaa on muokattu pitkäaikaissäilytyksen osalta. Hyväksyttiin muokattu työsuunnitelma.

Heikki Helin esitteli PAS-yhteistyöryhmän toimintaa. Esityksessä taustoitettiin aluksi pitkäaikaissäilytystä: mitä ovat säilyttämisen palvelut ja millainen on PAS-ratkaisu? CSC tuottaa opetus- ja kulttuuriministeriölle keskitettyjä PAS-palveluita. Keskeisenä tavoitteena on mahdollistaa digitaalisten aineistojen käyttökelpoisuus ja eheys pitkällä aikavälillä, jopa

28.3.2019

vuosisatojen ajan. Pitkäaikaissäilytyksestä sovitaan opetus- ja kulttuuriministeriön, CSC:n sekä aineistot omistavan organisaation välisillä sopimuksilla. Uusia PAS-sopimuksia solmitaan vasta kun GDPR yms. vaikutukset on voitu huomioida uusille sopimuksille. Kansainvälisestikin vertaillen Suomen PAS-järjestelmä on harvinainen. Aineistot eivät ole kaupallisten toimijoiden käsissä vaan järjestelmämme on täysin OKM:n rahoittama.

PAS-yhteistyöryhmää ovat edeltäneet PAS-jaosto, tukiryhmä ja yhteistyöryhmä, jotka ovat toimineet KDK-kokonaisarkkitehtuurityön rinnalla. Nykyisin PAS-yhteistyöryhmä on epävirallinen työryhmä, jossa ei ole nimitettyjä jäseniä. Ryhmällä on sähköpostilista ja kokouksia 3-4 kertaa vuodessa. Ryhmässä mm. vahvistetaan noudatettavat PAS-määrittelyt.

PAS-yhteistyöryhmän wiki, jossa julkaistaan pöytäkirjoja ja jatkossa myös luonnoksia:
<https://wiki.eduuni.fi/pages/viewpage.action?pageId=68256758>.

Viralliset julkaisut löytyvät vain www.digitalpreservation.fi -sivustolta.

Keskusteltiin erilaisista aineistoista ja niiden käsittelystä pitkäaikaissäilytyksessä, sekä käytiin läpi pitkäaikaissäilytyksessä olevan aineiston käytön mahdollisuuksia. Nykyisellään organisaation on mahdollista saada tarvittaessa käyttöönsä sama kokonaisuus, jonka se on siirtänyt palveluun säilytettäväksi. Erilaiset tiedostojen jakelukappaleet tuotetaan siten kunkin organisaation toimesta. Tulevaisuudessa mahdollisuudet voivat monipuolistua.

Juha Hakala esitteli pitkäaikaissäilytyksessä käytettäviä metatietoja. Metatiedot voidaan jakaa kolmeen ryhmään: kuvaileva, hallinnollinen ja rakenteinen. Pitkäaikaissäilytyksessä tarvitaan näitä kaikkia. Kuvailevaa metadataa ovat esimerkiksi tekijä, nimeke, tunniste, aihe ja moni muu sellainen tietoelementti, jota luetteloidessa tallennetaan. Teknistä metadataa ovat tiedoston koko, tiedostomuoto ja -versio sekä muut e-aineiston käsittelyssä tarvittavat tiedot, kuten vaikkapa still-kuvaa luotaessa käytetyt skannerin asetukset. Pitkäaikaissäilytyksen metadatalla kuvataan kaikki säilytyksen aikana tehdyt toimet, kuten tiedostojen kopiointi tai migraatio muuhun tiedostomuotoon. Migraatiosta voidaan tallentaa ajankohta, toimijat (ohjelmisto ja henkilö, joka toimenpiteestä vastasi) sekä migraation aiheuttamat muutokset tiedoston sisältöön ja ulkoasuun.

Pitkäaikaissäilytys on prosessi, joka alkaa kun dokumentti luodaan – ei vasta PAS-järjestelmään siirryttäessä. Jotta PAS-palveluun tulevien tiedostojen eheys voitaisiin varmistaa, pitäisi tiedostoille laskea tarkistussummat heti sen jälkeen, kun dokumentit on tallennettu julkaisuarkistoon. Yhteistyö on elintärkeää suositusten laatimisessa ja käyttöön saattamisessa. Esimerkiksi tiedostojen validointi pitäisi yhdenmukaistaa PAS-palvelun vaatimusten mukaiseksi jotta aineistojen siirto PAS-palveluun saataisiin mahdollisimman luotettavaksi.

Pitkäaikaissäilytyksessä tukeudutaan Pitkäaikaissäilytyksen viitemalliin eli OAIS-malliin, joka on myös suomennettu ja julkaistu SFS-standardina. Se määrittelee alan keskeiset käsitteet sekä PAS-toiminnan viitekehyksen toimijoinen ja palveluineen.

3. Paikkatieto-ryhmän tilannekatsaus

Paikkatietoryhmä kokoontui aloituskokoukseensa etäyhteydellä 15.3.2019. Ryhmässä ovat mukana Matias Frosterus, Kansalliskirjasto (Finto); Mikko Lappalainen, Kansalliskirjasto (Finto); Marja-Liisa Seppälä, Kansalliskirjasto; Miikka Haimila, Museovirasto; Sampsa Heinonen, Museoliitto; Miia Herrala, Kansallisarkisto ja Kenneth Ahlfors, Kansallisarkisto. Lisäksi pyydetään mukaan työskentelyyn Wikimedia Finlandin ja Finnan edustajat.

28.3.2019

Historiallisen paikkatiedon tietomallipohjan tekeminen aloitetaan 5.4. työkokouksessa. Työssä keskitytään ensisijaisesti historiallisen paikkatiedon mallintamiseen siten, että voidaan huomioida kuitenkin käytettävissä olevat muut paikan kuvailutiedon tuotokset. Lisäksi käydään läpi aiempien paikkatietoselvitysten materiaali. Syksyllä pidetään Historialliset paikat kick-off -tilaisuus/työpaja jossa keskustellaan laajemmin ryhmän tekemän pohjatyön perusteella. Työpaja vaatii vielä tarkempaa suunnittelua.

Paikkatietoryhmässä koettiin, että konkreettiseen työhön ryhtymiseen ei ole esteitä, sillä tarpeet on jo aiemmin määritelty KDK-tietoarkkitehtuuriryhmän selvityksessä.

4. Muut asiat

– **YSAsta YSOon siirtymä** (Matias Frosterus)

YSAsta YSOon siirtymä toteutuu kesän aikana. Siirtymä on keskeisin kirjastoille, mutta sillä on laajemmat vaikutukset koko KAM-sektorille. Siirtymässä konvertoidaan olemassa olevat ysa-termit ja jatkossa kuvaillaan yso-käsitteitä käyttäen. Sanastoja on muutoksessa mukana enemmän kuin vain ysa-yso: mm. allärs, suomalainen lajityyppi- ja muotosanasto (slm), suomalainen esityskokoonpanosanasto (Seko.) Ysa jäädytetään, samoin allärs ja ne löytyvät edelleen Fintosta, mutta eivät päivity enää.

Sanastot ovat vapaassa käytössä, joten Kansalliskirjastossa ei ole tietoa, missä kaikkialla sanastoja käytetään. Kansalliskirjastossa laaditaan tiedote, jota toivotaan myös Tietoarkkitehtuuriryhmän jäsenten jakavan edelleen omille yhteistyötahoilleen. Digime-tiedotusta on tarkoitus myös hyödyntää. Sovittiin, että Kansalliskirjaston laatima tiedote jaetaan tietoarkkitehtuuriryhmän jäsenille.

– **Ajan esittäminen kuvailussa** (Mikko Lappalainen)

Millä tavoin ja mitä standardeja hyödyntäen eri sektoreilla viitataan aikaan tai ajankohtiin, esim. kirjan julkaisuaika, esineen luomisaika? Aikaan viitataan monella tavalla kaikilla KAM-sektoreilla, joten voisi olla luontevaa miettiä asiaa tietoarkkitehtuuriryhmässä.

Kirjastoilla ollut eri aikoina eri ohjeistuksia ja erilaisia ajan merkitsemisen tapoja on käytetty hyvin heterogeenisesti. Tilanne on samanlainen, ellei jopa huonompi museoissa. Arkistoissa ollaan merkitsemistavoissa lähempänä standardia. Todettiin, että kokonaisuus pitäisi saada standardisoitua kautta KAM-sektorin. ISO 8601 -standardi on päivitetty helmikuussa 2019 ja mukana on käytännössä kaikki kirjastojen tarvitsemat ominaisuudet, jotka on aiemmin kuvattu Kongressin kirjaston EDTF (Extended Date/Time Format) -profiilissa. Standardin osa yksi kuvaa yleiset piirteet, osa 2 laajennukset. Käytännössä useimmat kuvailun tarpeet kytetään hoitamaan osassa 1 kuvailtujen ominaisuuksien avulla. Osassa 2 on myös sellaisia piirteitä, joita yksikään muistiorganisaatio tuskin tarvitsee.

Keskusteltiin eri vaihtoehdoista tilanteesta. Todettiin, että parasta olisi, jos laadittaisiin ISO 8601:n kansallinen soveltamisprofiili. Profiili voisi kattaa esimerkiksi osan yksi kokonaisuudessaan, ja lisäksi osasta kaksi piirteet, joita voi tai pitää voida tukea. Tämä pitäisi viestiä niin, että kaikki järjestelmät pystyvät tukemaan profiilin pakollisten ominaisuuksien käyttöä. Poikkeamat voitaisiin myös määritellä. Tällainen poikkeama voisi olla esimerkiksi ajanjakso, joka standardin mukaan määritellään kauttaviivalla (1933/1992), mutta kuvailussa käytetään yleisesti tavuviivaa (1933-1992). Profiili olisi mahdollista myös virallistaa SFS:n tietohuoltokomitean kautta. Näin muutkin yhteisöt saisivat tiedon profiilista ja voisivat hyödyntää sitä.

28.3.2019

Päätettiin ottaa asia tietoarkkitehtuuriryhmässä syksyn työlistalle. Tavoitteena on saada työ melko valmiiksi vuoden loppuun mennessä, sillä myös RDA-soveltamisprofiili tulisi saada aikamääreidenkin osalta valmiiksi. Juha Hakala lupasi toimia yhteyshenkilönä SFS:ään. KAM-kuvailuryhmä voi toimia sektorien käytänteiden ja tarpeiden kerääjänä. Asian käsittelyyn varataan erillinen kokous, johon halukkaat voivat osallistua. KAM-kuvailuryhmä käsittelee RDA-sovellusprofiilia touko-kesäkuun kokouksessaan. Pidetään asiasta katsaus seuraavassa Digime-tietoarkkitehtuuriryhmän kokouksessa ja sovitaan sitten työn tarkemmasta organisoimisesta.

- **Kirjastoverkkopäivien (22.–24.10.2019) suunnittelu** on aloitettu. Onko KAM-sektorilta toiveita ohjelmaan?

Kirjastoverkkopäivien ohjelmassa on puolenpäivän ja päivän mittaisia työpajoja (ensimmäinen päivä) sekä esityksiä. Aihe-ehdotuksia työpajoiksi tai esityksiksi voi lähettää Matiakselle (matias-frosterus@helsinki.fi) tai Ninalle (nina.hyvonen@helsinki.fi) huhtikuun aikana.

- **Tiedoksi: Näyttörajoitettuihin aineistoihin** liittyvissä huomioissa ottakaa yhteys Piia Naukkariseen (piia.naukkarinen@helsinki.fi)

5. Kokouksen päättäminen

Puheenjohtaja päätti kokouksen klo 12.10.