

2022-11-25 - Julkaisuarkistopalveluiden asiantuntijaryhmän kokous 4/2022

Aika: 25.11.2022 klo 9:15-11:00

Paikka: Zoom

Läsnä

Ryhmän jäsenet

- Byholm, Leena (Luonnonvarakeskus / Jukuri)
- Kallio, Merja (Vaasan yliopisto / Osuva) - puheenjohtaja
- Kejonen, Eetu (Åbo akademi / Doria)
- Lehtilä, Tapani (Tampereen yliopisto / Trepo)
- Liikala, Minna (Terveystieteiden ja hyvinvoinnin laitos / Julkari)
- Lindström, Susanna (Suomen Pankki / Kaisu)
- Maisalmi, Outi (Väylävirasto / Doria)
- Marjamaa, Minna (Laurea-ammattikorkeakoulu / Theseus)
- Myllyharju, Mari (Valtioneuvoston kanslia / Valto)
- Niskanen, Anita (Valtion taloudellinen tutkimuskeskus / Doria)
- Nokelainen, Ilkka (LUT-yliopisto / LUTPub)
- Nurmi, Erkki (Taideyliopisto / Taju)
- Nyqvist Antti (Haaga-Helia ammattikorkeakoulu / Theseus)
- Pennanen, Mikko (Turun yliopisto / UTUPub)
- Raja-Hanhela, Toni (Lapin yliopisto / Lauda)
- Tolonen, Tiina (Oulun ammattikorkeakoulu / Theseus)

Kansalliskirjaston asiantuntijat

- Hahl, Päivi
- Ilva, Jyrki
- Kuuskoski, Reeta - sihteeri
- Lindholm, Erik (Vapaakappalepalvelut)
- Moubarik, Anis
- Mujunen Veera (Finna)
- Niininen, Satu (Kansallisbibliografiapalvelut)
- Rajakangas, Jukka-Pekka (Vapaakappalepalvelut)
- Viita, Samu

Muistio

1. Kokouksen avaus / Merja Kallio
 - a. Kokous avattiin klo 9:17.
 - b. Asiantuntijaryhmän uutena jäsenenä aloittaa Suomen Pankin Susanna Lindström
 - i. Susanna on julkaisuarkiston pääkäyttäjä. Suomen Pankin kirjasto on pieni organisaatio, jossa työskentelee kolme henkilöä, ja toimenkuva on kaikilla laaja.
 - c. Miika Oksman on aloittanut tietojärjestelmäasiantuntijana Kansalliskirjaston julkaisemisen palveluissa
 - i. Miika aloitti työt julkaisemisen palveluissa n. kuukausi sitten. Hän on aiemmin työskennellyt mm. Aalto-yliopistossa ja yksityisellä sektorilla.
 - d. Ryhmän jäsenten esittäytymiskierros
2. Suomen Pankin ja Finanssivalvonnan julkaisuarkisto [Kaisu](#) avattiin 22.11.22 / Susanna Lindström
 - a. [Esityksen diat](#)
 - i. Kaisu avattiin tiistaina 22.11.22
 - ii. Julkaisuarkisto on tärkeä tekijä julkaisuprosessissa: peruseriaate on, että julkaisu tallennetaan vain kerran, ja integraatioiden kautta se leviää mm. organisaation verkkosivuille.
 - iii. Ennen Kaisua Suomen Pankin julkaisuarkisto oli Helsingin yliopiston Heldassa, jossa se sijaitti loppuvuodesta 2014 alkaen.
 1. Heldassa on auki Kaisun rinnalla vielä parisen viikkoa, jotta mm. integraatioiden toiminta saadaan varmistettua.
 2. Tähänastisessa julkaisuarkistossa on ollut n. puoli miljoonaa latausta vuodessa. Suurin osa kävijöistä tulee Googlen kautta.
 3. Eniten ladattu julkaisu on [analyysi bitcoinista](#), sitä on ladattu lähes 70 000 kertaa.
 - iv. Keskustelut julkaisuarkiston siirtymisestä Kansalliskirjaston ylläpitoon aloitettiin n. vuosi sitten, sopimus allekirjoitettiin alkuvuonna 2022.
 1. Toukokuussa alkoi ympäristön pystytys, ja ennen juhannusta tehtiin ensimmäiset migraatiot Heldasta.
 - a. Hektisin vaihe alkoi syyskuun alusta, integraatiot olivat projektissa keskeisessä asemassa.
 - i. Nostot Suomen Pankin ja FIVAn verkkosivuille ovat hyvin olennaisia.
 - ii. Finna-integraatio on todettu toimivaksi Kaisusta.
 - b. Projektin kuluessa Kansalliskirjaston kanssa yhteistyö on sujunut loistavasti. Työtä on ollut paljon, mutta asiat ovat edistyneet ja kysymyksiin on vastattu nopeasti.
 - b. Merja Kallio kommentoi, että Suomen Pankin Heldassa olevat julkaisut ovat tulleet vastaan Vaasassa Osuvaan tallennettujen julkaisujen lähdeluetteloissa.

- c. Veera Mujunen totesi, että Finna toimii alustansiirron jälkeen, ja kansikuvatkin näkyvät oikein. Veera kertoi, että Finnassa työskennellään saatavuustietojen parissa. Tarkoitus on, että Finnan käyttöliittymässä olisi selvästi merkitty tieto siitä, mitkä aineistot ovat luettavissa vapaasti verkossa.
3. PAS-tilannekatsaus / Jukka-Pekka Rajakangas
- Paketointi ja lähettäminen CSC:lle onnistuu. Viime aikoina on testattu pdf- ja epub-formaatissa tallennettujen evapaakappaleiden luovutusten lähettämistä. Tuotantolähetyksiä on tehty 600 kpl, joista 500 meni läpi. Epäonnistuneissa olleita virheitä selvitetään. Mukana on kummallisia tiedostomuotoja ja pdf-tiedostoja, joiden validointi CSC:n validaattorilla ei ole onnistunut.
 - Joulukuussa aloitetaan Kansalliskirjaston Mikkelin toimipisteen digitoimien aineiston lähetysten testaaminen. Mikkelin digitointiyksikkö tuottaa METS-paketti, vapaakappaleyksikkö lähettää ne keskitetysti PAS-järjestelmään. Näin saadaan tarvittavat tiedot Kansalliskirjaston PAS-tietokantaan.
 - Tapani Lehtilä tiedusteli, millaiset versiot eivät mene validoinnista läpi. Tällaisia ovat mm. salatut tiedostot, joita testilähetyksissä on muutama mukana. Lisäksi epäonnistuneiden joukossa oli vanhempia epub- ja pdf-versioita, joita ei tällä hetkellä tueta. Jälkimmäisten kanssa ei vielä ole päätetty, miten niiden kanssa toimitaan. Niille tehdään ehkä migraatio, mutta toistaiseksi ne jäävät odottamaan käsittelyjonoon päätöstä siitä, minkälaiseen ratkaisuun päädytään..
4. Ajankohtaista Finna-asiaa / Veera Mujunen
- Esiityksen diat**
 - Theseuksen aineistotyyppejä päivitetty
 - Uusi äänikirja-aineistotyyppi tulossa
 - Jos aineistossasi on äänikirjoja, ota yhteyttä Finnaan
 - Pääsyrjaustietojen tunnistamista on parannettu haun rajauksen osalta, mutta tietuesivulla tieto ei näy selkeästi
 - Embargomerkinnät ovat hankalia Finnassa, merkitsemiskäytännöt metadatatassa vaihtelevat
 - Oikopolku tutkittuun tietoon -sivu julkaisu vuosi sitten
 - Aineistolähteinä julkaisuarkistot ja journal.fi
 - Poimintoja ajankohtaisista tutkimusjulkaisuista -nostokaruselliin voi ehdottaa julkaisuja, [ohje tähän on Finnan asiakaswikissä](#)
 - Finnaan on tulossa vielä uusia julkaisuarkistoja
 - Veera on ehtiessään yhteydessä kiinnostuksensa osoittaneihin organisaatioihin, mutta asiaa voi tarvittaessa kiirehtiä ottamalla yhteyttä Veeraan
 - Kokotekstihaun tutkiminen jatkuu, tällä hetkellä Finnan haku kohdistuu vain julkaisuarkistojen metatietoon
 - Kokotekstihaun toteuttaminen vaatii käytettävyyssuunnittelua ja lisää palvelinkapasiteettia
 - Alkuvuodesta varmistuu, voidaanko projektia jatkaa
 - Ennen toteuttamista asiasta sovitaan asiakkaiden kanssa
5. Julkaisuarkistopalveluiden ajankohtaisia asioita / Jyrki Ilva ja Samu Viita
- DSpace 7-päivitysten suunnittelu
 - Lyrasis järjesti marraskuussa [kolme DSpace 7.4 -webinaaria](#)
 - Ensimmäinen tilaisuus kertoo DSpace 7 -versiosta yleisemmällä tasolla, kaksi jälkimmäistä ovat luonteeltaan teknisempiä
 - Kansalliskirjastolla alkuvuosi 2023 menee päivitysten suunnitteluun, päivityskierros alkaa todennäköisesti ensi syksynä ja kestää pitkälle vuoden 2024 puolelle. Samalla tehdään muitakin teknisiä uudistuksia julkaisuarkistoihin.
 - DSpace 7 -käyttöönottoja on maailmalla tehty vielä suhteellisen vähän
 - Suuret organisaatiot eivät ole vielä siirtyneet käyttämään DSpace 7-versiota
 - Kansalliskirjastoa kiinnostaa mm. järjestelmän skaalautuminen isoihin käyttömääriin
 - Päivitysten suunnittelun tavoitteena mm. palvelun kehittämisen ja ylläpidon virtaviivaistaminen
 - On määritettävä hyvä perustuote, jotta asiakaskohtaista räätälöintiä tarvitaan mahdollisimman vähän
 - Modulaarisuus mahdollistaa DSpace:n pilkkomisen
 - Esim. SOLR-hakumoottori on DSpace 7 -versiossa irrallinen kokonaisuus
 - Tietokanta on erotettavissa
 - AssetStore eli tiedostojen tallennuspaikka saadaan erilleen
 - Kokemusta toimintojen hajauttamisesta on jo Theseuksesta, joka koostuu viiden palvelimen klusterista. Erilaisia suunnitelmia arkkitehtuurille on, mutta työ on vielä kesken.
 - Päivitykset alkavat näillä näkymin ensi syksynä
 - Arkistojen päivitysjärjestys ei ole vielä tiedossa
 - Asiaan palataan myöhemmissä kokouksissa
 - Miika Oksmanin työpanos tulee olemaan suuri apu päivityskierroksessa
- Palvelunestohyökkäysten torjunta / Päivi Hahl
 - Yhteydenottolomake on poistettu julkaisuarkistoista
 - Arkistojen yhteydenotto-sivulla (contact) on nyt vain sähköpostiosoite, ei enää linkkiä palautealomakkeelle
 - Kun asiakas käytti lomaketta, viesti tuli Kansalliskirjaston osoitteesta, mutta nyt lähettäjänä on oikea käyttäjä
 - Organisaation roskapostisuodatus toimii todennäköisesti paremmin
 - Muitakin ratkaisuja selvittävänä
 - Pidemmän tähtäimen suunnitelmat toteutuvat todennäköisesti versionvaihdon yhteydessä
 - Vaihtoehtoja on useita, vaativat testausta
 - Kansalliskirjastolla on pohdittu mahdollisuutta ottaa käyttöön esim. [Cloudflare](#)
 - Laajasti käytetty kaupallinen palvelu, joka toimii proxyinä arkiston edustalla
 - Tunnistaa haitallisen liikenteen ja epäilyttävät IP-osoitealueet ja estää niiden pääsyn palvelimelle
 - Julkaisuarkistopalvelin saisi tällöin vain oikeat kyselyt
 - Palvelu on maksullinen
 - Organisaatiotasolla on mietittävä hallintamalli eli se, miten kussakin arkistossa Cloudflare otettaisiin käyttöön:
 - Antaisiko asiakas Kansalliskirjastolle luvan tehdä käyttöönoton puolestaan, jolloin domainin hallinta on edelleen asiakkaalla ja vaatii asiakkaan pientä IT-panosta (KK suosittaa).
 - Rekisteröityisikö asiakas itse Cloudflareen ja antaisi Kansalliskirjastolle tarvittavat admin-oikeudet
 - Se, onko kyseessä subdomain tai onko julkaisuarkistolla oma domain, vaikuttaa mahdollisuuksiin
 - Tapani Lehtilä tiedusteli asiakasorganisaatiolle tulevia kustannuksia
 - Kustannuksiin vaikuttaa valittu hallintamalli ja palvelutaso
 - Asia on vielä selvittelyn alla, mutta arkistoa kohden kyse ei välttämättä ole kovin suurista summista
 - Yksityiskohtiin palataan tarvittaessa myöhemmin
- Saavutettavuusselosteisiin tehty pieniä päivityksiä / Reeta Kuuskoski

- i. Päivityksessä korjattiin ruotsinkielisten selosteiden typo ja korvattiin englanninkielisissä selosteissa olevat AVIn suomen- ja ruotsinkielisille sivuille vievät linkit uudelle englanninkieliselle AVIn sivulle vievällä linkillä.
 - ii. Koska muutokset eivät vaikuttaneet selosteen varsinaiseen sisältöön, päivityspäivämäärää ei muutettu.
 - d. Efecte-tikettijärjestelmä käyttöönotto suunnitteilla ensi vuoden puolella / Reeta Kuuskoski
 - i. Nykyinen palveluosoite doria-oa(at)helsinki.fi korvautuu toisella.
 - ii. Tämä on vasta etukätestiedotusta, työ on vielä suunnitteluvaiheessa.
 - iii. Asiasta ilmoitetaan, kun osoite muuttuu
- 6. Muut asiat
 - a. Kirjasto verkkopalveluiden asiakaskyselyn tulokset / Jyrki Ilva
 - i. Julkaisuarkistopalveluille positiivista palautetta
 - 1. Ks. [yhteenveto tuloksista](#)
 - 2. Vastauksia oli vähän vähemmän kuin edellisellä kerralla. Vastauksia saatiin yhteensä 197, ja niistä 37 koski julkaisuarkistopalveluita. Kysymykset olivat samat kaikille Kansalliskirjaston palveluille.
 - a. Kansalliskirjaston työntekijöiden asiantuntemusta kiiteltiin, palveluiden käytettävyys sai vähän huonompia arvosanoja.
 - b. Kustannussäästöjä pidettiin merkittävänä etuna
 - c. Tallennus- ja kuvailutyökalut saivat heikoimmat arvot, samoin haulle kritiikkiä
 - ii. Erkki Nurmi totesi, että kyselyyn vastatessa oli ongelmana vastattavien kysymysten valinta
 - 1. Vaikka Taideyliopistossa on käytössä Melinda, ei sitä voi arvioida, jos sen kanssa ei tee töitä. Olisi hyvä ilmaista selvemmin se, että kannattaa vastata vain niiden palvelujen kysymyksiin, joista itse tietää jotakin.
 - 2. Kansalliskirjasto vie asiaa eteenpäin, toivotaan, että seuraava kysely on tältä osin selkeämpi.
 - b. Opinnäytteiden avoimuutta koskeva linjaus / Merja Kallio
 - i. Linjausta valmisteleva työryhmä aloittanut työnsä
 - 1. Työryhmässä on mukana SYLin, UNIFIn, FUNin, SAMOKin, AMKIT-konsortion, Professoriliiton ja ARENEn edustajia. Asiantuntijaryhmän jäsenistä työssä on mukana Merja Kallio, Minna Marjamaa, Eetu Kejonen ja Jyrki Ilva
 - 2. Linjauksen laatimiseen on aikaa noin vuosi, ensi kevääksi pitäisi saada jo luonnos kommentoitavaksi. Työryhmä kokoontuu kuukausittain.
 - 3. Ensimmäisenä asiana ryhmässä on määritelty termejä: mikä on avoin, mikä on julkinen, mikä on opinnäyte? Lisäksi on keskusteltu tausta-aineistosta eli datasta ja sen avoimeksi asettamisesta.
 - 4. Työryhmän työn edistymisestä kerrotaan jatkossakin asiantuntijaryhmässä
 - 5. [Työryhmän marraskuun kokouksen muistio](#) on luettavissa.
 - c. Uusia DSpace-jäsenorganisaatioita Suomesta (FinELib/SCOSS) / Jyrki Ilva
 - i. Ks. lista organisaatioista [DSpace-wikissä](#)
 - 1. FinELib-konsortion edustajaksi DSpace Leadership Groupiin Jussi Piipponen HuLibistä
 - 2. Jyrki Ilva jatkaa Kansalliskirjaston edustajana
 - 3. Molemmat pyrkivät osallistumaan toimintaan aktiivisesti ja tekemään yhteistyötä
 - 4. Ensimmäinen uuden ryhmän kokous on [keskiviikkona 30.11.](#)
 - d. [OKM:n konsernin kehittämishanke](#) / Jyrki Ilva
 - i. Tavoitteena on vähentää OKM:n alaisten virastojen määrää
 - ii. Arkistot ja kirjastot -virastokokonaisuuden koostumukseksi kaavailtiin aluksi Kansallisarkistoa, Varastokirjastoa, Celiaa ja Kotuksen kirjastopalveluita
 - 1. Myös Kansalliskirjaston asema nostettiin loppusuoralla mukaan selvitystyöhön
 - 2. Lokakuussa julkaistu raportti [on luettavissa Valtossa](#)
 - 3. Prosessin lopputuloksesta ei vielä tietoa
 - iii. Kansalliskirjasto on nyt Helsingin yliopiston erillislaitos
 - 1. Tällä hetkellä odotetaan yliopiston tekemää selvitystä asiasta
 - 2. Irtautuminen yliopistosta ja yhdistyminen muiden virastojen kanssa olisi Kansalliskirjastolle radikaali muutos
 - 3. Jos muutos toteutuu, se vaikuttaa todennäköisesti myös muille organisaatioille tarjottaviin palveluihin, ml. julkaisuarkistopalvelut
 - e. [Edellisessä kokouksessa](#) käsitellyn esteettömyysseminaarin jatkokeskustelu lain tultua voimaan siirtyy ensi vuodelle / Reeta Kuuskoski
 - i. [Laki ei ole vielääkään tullut voimaan](#) (23.11.)
 - ii. AVIn Eetu Komi tai hänen kollegansa puhumaan johonkin ensi vuoden kokoukseen, mutta vielä ei tiedetä, mihin
 - Eduskuntavaalit tulossa keväällä, lain voimaantulosta ei ole tietoa
- 7. Seuraavat kokoukset
 - a. Päätettiin pitää ensi keväänä vain yksi kokous, joka toteutetaan kasvotusten ns. live-kokouksena
 - i. hybridiosallistumismahdollisuudet ovat KK:n tiloista melko hyvät
 - ii. Reeta Kuuskoski tekee Doodlen / Framadaten päivän valinnaksi
 - 1. Ajankohta on huhtikuun loppupuolella
 - 2. Aloitusaika aikaisintaan klo 10
 - 3. Olisi hyvä, jos päivän tienoilla olisi jotain muutakin eikä pelkkä kokous