

Theseus-julkaisuarkiston ohjausryhmän kokous 4 / 2015

Aika: Maanantai 27. 4. 2015.

Läsnä: Teemu Makkonen, puheenjohtaja
Kaisu Kinnunen, amk-opettajien edustaja
Tommi Pälli, amk-kirjastovirkailijoiden edustaja
Hanna Saario, amk-kirjastonjohtajien edustaja
Nina Hyvönen, Kansalliskirjaston kirjastoverkkopalvelut

Asiantuntijat:
Anna-Liisa Holmström, AMKIT-konsortio
Minna Marjamaa, AMKIT-konsortio
Tiina Tolonen, AMKIT-konsortio
Samu Viita, Kansalliskirjaston kirjastoverkkopalvelut (sihteeri)

1. Kokouksen avaus

Avattiin kokous ja todettiin läsnäolijat. Riitta-Liisa Karjalainen luopui paikastaan ohjausryhmässä ja Hanna Saario on nyt uusi amk-kirjastonjohtajien edustaja. Saario on kirjasto- ja tietopalvelupäällikkö Diakonia-ammattikorkeakoulusta.

2. Edellisen kokouksen muistio

Teemu Makkonen esitteli edellisen kokouksen muistion ja se hyväksyttiin.

3. Järjestäytyminen

Puheenjohtajana toimi Teemu Makkonen, kokoukselle valittiin sihteeriksi Samu Viita.

4. Theseuksen kehittäminen jatkossa

SYLI-syöttöjärjestelmä

Åbo Akademi pilotoi gradujen syöttämistä opiskelijavetoisesti Doriassa. Pilotin esiintuomien uusien tarpeiden ja havaittujen bugien korjaaminen on hidastanut käyttöönottoa laajemmin. Myös resurssipula hidastaa käyttöönottoja, prioriteettilistalla SYLIn kehittäminen tulee varsinaisten julkaisuarkistopalveluiden ylläpidon ja kehittämisen jälkeen.

Keskusteltiin siitä, mitä amk:t odottavat SYLI:ltä nykyiseen syöttöprosessiin verrattuna. Tärkeimmiksi kohdiksi mainittiin, että SYLI:n avulla jo tarkistusvaiheessa voidaan muokata metadataa. Toinen pääsyy on streamauspalvelu, jota nykyinen syöttöprosessi ei tue. SYLIin toivotaan helppokäyttöistä toteutusta tukemaan streamattavan materiaalin syöttöä ja tarkistamista.

Streamattavien aineistojen osalta haasteita on myös oikeudellisten seikkojen kanssa.

Tekijänoikeusselvitys Theseuksen streamattavan aineiston osalta on tehty, mutta asiaan liittyy vielä käytännössä muita epäselviä asioita, esimerkiksi tarvitseeko kaikilta videolla esiintyviltä pyytää lupa julkaisuun, esim. viittomakielisen materiaalin osalta.

Puheeksi tuli myös kustannukset mahdollisesta Sylin käyttöstä Theseuksessa. Kansalliskirjaston lähtökohtana on tarjota SYLI ja streamauspalvelu osana julkaisuarkistopalvelua, eikä siitä veloiteta erikseen. Ainoastaan isommat muutokset itse palvelun toimintaan voivat aiheuttaa kustannuksista, mutta tekniikan vaihdos ei vaikuta itsessään hintaan.

Samu Viita ehdotti, että Syliä lähdetään pilotoimaan yhden tai kahden amk:n voimin ja tähän kokemukseen perustuen yritetään räätälöidä SYLI Theseukselle sopivaksi. Haasteena Sylin laajempaan käyttöön ottoon Theseuksessa on mm. se, että Theseuksen nykyinen syöttölomake ja syöttöprosessikonaisuus on räätälöity pitkälle Theseuksen tarpeisiin ja jotkin räätälöidyt ominaisuudet voi olla vaikeita toteuttaa SYLI:n avulla ainakaan samalla tavalla. Tämän vuoksi on hyvä, että asiaa pilotoidaan ensin huolella ja vaatimuksia tarkennetaan kerätyn kokemuksen perusteella. Myös ohjeiden laatimista ja käyttöliittymän kielikäännöksiä on tiedossa runsaasti ja niiden määrittely ottaa aikansa.

Theseus ja tutkimusdata

Pohdittiin myös voisiko Theseusta käyttää tutkimusdatan tallentamiseen. Todettiin, että periaatteessa pystyy, mutta asia riippuu siitä, millaisia vaatimuksia se edellyttää ja että onko järkevää tallettaa lähinnä opinnäytetietokantana profiloituneeseen palveluun luonteeltaan hyvin erilaista materiaalia. Nina Hyvönen nosti esiin, että voisi olla järkevämpää käyttää erityisesti tutkimusdataan keskittyvää palvelua tähän tarkoitukseen. Todettiin myös, että tutkimusmateriaalissa on myös paljon juridisia seikkoja huomioitavina, esimerkiksi se, että arkaluontoista materiaalia, kuten henkilötietoja sisältävää dataa ei päädy julkiseen levitykseen.

Theseus ja rinnakkaistallentaminen

Keskusteltiin Theseuksesta rinnakkaistallennuspaikkana. Rinnakkaistallentaminen on ollut mahdollista tekniikan osalta Theseuksessa jo vuodesta 2010 saakka, mutta tallennusten määrä ei ole lähtenyt odotetusti julkaisujen osalta liikkeelle. Joillakin amkeilla löytyy Theseuksesta julkaisuja, mutta volyymi on melko pieni siihen nähden, kuinka paljon esimerkiksi ammattikorkeakoulut tuottavat artikkeleita alan ammattilehtiin.

Nina Hyvönen mainitsi, että Kansalliskirjastossa on alkamassa avoin tiede ja tutkimus -hankkeen (ATT) osahanke 1.9. Siinä on tarkoitus edistää mm. rinnakkaistallentamista ja kartoittaa OA-kirjoittajamaksujen keräämistä Suomessa. Nina suositteli seuraamaan hankkeen edistymistä myöhemmin avattavalta Kansalliskirjaston ATT-hankeen WWW-sivustolta.

Theseuksen rinnakkaistallennuslomakkeelle haluttaisiin embargo-ajan määrittämismahdollisuus. Sovittiin että Theseus-asiantuntijat hoitavat tämän asian keskenään kuntoon.

Rinnakkaistallennukseen liittyen pohdittiin, millaisia rinnakkaisjulkaisuja Theseukseen viedään ja onko opinnäytteitä etsiville opiskelijoille hämmentävää, jos Theseuksesta löytyykin jatkossa huomattavan paljon erilaisia amk-julkaisuja. Päädyttiin siihen, että tilannetta pystytään

selkeyttämään kuvailutietojen ja kokoelmajaottelujen avulla. Ohjausryhmä ei päätä sitä, mikä on validi rinnakkaistalenne vaan se arvioidaan ammattikorkeakouluissa.

Päätettiin, että tiedotetaan ja markkinoidaan jälleen kerran kirjastoille mahdollisuutta tallentaa artikkeleita ja muita henkikökunnan tuottamia julkaisuja Theseukseen. Ehdotettiin myös, että rukkaset voisivat tehdä wikiin tietoisuutta rinnakkaistalennuksen asioista, mm. SHERPA/RoMEO:sta yms. aiheeseen liittyvistä työkaluista.

Theseuksen tilastointitarpeet

Keskustelua ovat herättäneet myös tilastot, joita kaipailtaisiin enemmän. Samu Viita pyysi toimittamaan listaa, mitä tilastoja tarvitaan. Rukkaset kyselevät kentältä tärkeimmistä tilastointitarpeista ja niiden perustella valitaan hyödyllisimmät ja toteuttamiskelpoisimmat. Anna-Liisa Holmström nosti esiin, että erityisesti kartuntatilastot olisivat tärkeitä. Kansalliskirjaston tekemään Simplestats-tilastointiin ollaan oltu tyytyväisiä ja sen JSON-laajennusta käytetään jo Oulun ammattikorkeakoulussa. JSON-laajennus mahdollistaa sen, että tilastoja voi kysellä asiakaspäädyn sivustolta JavaScriptin avulla, jolloin haluttuja tilastoja voi koostaa omille sivuilleen.

Tekniset kuormitusongelmat

Theseuksessa on esiintynyt viime aikoina häiriöitä, joiden kesto on ollut maksimissaan noin viisi minuttia. Tällaisia häiriöitä on voinut olla yhdestä kahteen päivän aikana. Kansalliskirjastossa selvitetään syytä paraikaa, tämänhetkiset epäillyt ovat Dspace 3 version mukanaan tuoma, kokonaan uusittu OAI-PMH toteutus, jossa epäillään esiintyvän muistivuotoa. Muistivuoto-ongelma on oireillut kunnolla vasta Theseuksen kuormituksen kasvaessa kevään ruuhkapiikin alkaessa.

5. Rekisteriseloste

Theseuksen rekisteriselostetta on muokattu ja muokkaamiselle on tarvetta edelleen. Kokouksessa kiinnitettiin huomioita seuraaviin sanamuotoihin nykyisessä tekstissä: Termi koulutuslinja on vanhentunut, sille etsitään OKM:n sivuilta oikea ja ajanmukainen termi, todennäköisesti "tutkinto-ohjelma". Termi valmistumisvuosi otetaan pois, samoin opinnäytetyön arvosana, sillä Theseukseen ei tallenneta arvosanatietoja. Myöskään käyttäjäsuhteen päättymistieto ei ole Theseuksessa, joten päätettiin ottaa sekin pois.

Rekisteriselosteen kohta 10. herätti keskustelua, sillä Theseuksessa ei ole toteutettu mahdollisuutta kieltää tietojen luovuttamisesta suoramarkkinointiin yms. Theseuksen tietoja ei kuitenkaan missään tapauksessa tällaisiin paikkoihin muutenkaan luovuteta, joten teksti on turha.

Teemu Makkonen nosti esiin, että rekisteriselosteessa puhuttaan vain opinnäytetyöstä, vaikka julkaisusana olisi parempi, sen sopiessa kaikkiin tilanteisiin. Päätettiin muuttaa opinnäytetyö-termi julkaisu-termiksi.

Kun suomenkielinen teksti on valmis, käännetetään tekstit myös ruotsiksi ja englanniksi.

Pohdittiin, laitetaanko maininta, että metadata on avoimesti hyödynnettävissä (ja esim. haravoitavissa ja siirrettävissä muihin järjestelmin). Samassa yhteydessä Theseuksen wikin FAQ-osioon päätettiin myös lisätä, että metadata ei ole tekijänoikeuden alaista. Metadatan tekijänoikeuksista heräsi keskustelua, sillä yliopistoissa on ollut epäselvyyksiä, saako metadataa siirtää esimerkiksi OAI-PMH -rajapinnalla muihin järjestelmiin. Todettiin, että Theseuksen metadataan ei kohdistu tekijänoikeudellisia ongelmia ja se on vapaasti siirrettävissä CC0-lisenssillä.

6. Muut asiat

Theseusta ei ole toistaiseksi päätetty kilpailuttaa uudestaan.

Kansalliskirjastolta on pyydetty tiedonhallinnan ohjausryhmässä, että se perustaa alaryhmän, joka keskittyy julkaisuarkistoihin. Perustettu ryhmä on nimeltään julkaisuarkistojen asiantuntijaryhmä ja siinä on mukana organisaatioita, joille Kansalliskirjasto tuottaa julkaisuarkistopalveluita. Ryhmän ensimmäinen tapaaminen oli 9.4.2015 ja tapaamisen muistio on luettavissa osoitteesta:

<https://www.kiwi.fi/display/JULPAL/2015-04-09>

Myöhemmin, kun ATT-hanke tulee ajankohtaisemmaksi, on ehkä tarvetta perustaa myös sitä varten oma ryhmänsä.

Amk-kentältä on tullut kyselyitä, miten menetellään kun suomalaisen ammattikorkeakoulun opiskelija tekee yhteisen opinnäytetyön jonkin toisen ammattikorkeakoulun kanssa? Tästä on ollut keskusteltua jo vuosia takaperin ja on päädytty siihen, että kunkin ammattikorkeakoulun oppilas tallentaa omaan ammattikorkeakoulunsa kokoelmaan yhteisesti tehdyn opinnäytetyön. Jos puolestaan opinnäyte tehdään yhteiseen koulutusohjelmaan, niin tallennus tapahtuu vain kyseisen koulutusohjelman kokoelmaan. Sama työ on siis ensimmäisessä skenaariossa useampaan kertaan tallennettuna. Tekijät tallentavat oman oppilaitoksensa ohjeiden mukaan, joten esim. kansilehti saattaa poiketa toisistaan.

Theseus-posteri hyväksyttiin kesäkuun LIBER2015 konferenssiin Lontooseen. Rahoitusta haetaan konsortiolta ja Suomen kirjastoseuralta, Anna-Liisa Holmström, Minna Marjamaa ja Tiina Tolonen ja ovat lähdössä Liberiin esittelemään posteria. Konsortio on tukemassa matkaa rahallisesti. Posteria esiteltiin kokouksessa.

7. Kokouksen päättäminen

Kokous päätettiin 14.29