

2020-09-30

Julkaisuarkistopalveluiden asiantuntijaryhmä, 30.9.2020

Aika ke 30.9.20 klo 14-16

Paikka Zoom

Läsnä

- Leena Byholm (Luonnonvarakeskus / Jukuri)
- Merja Kallio (Vaasan yliopisto / Osuva)
- Eetu Kejonen (Åbo akademi / Doria)
- Tapani Lehtilä (Tampereen yliopisto / Trepo)
- Outi Maisalmi (Väylävirasto / Doria)
- Minna Marjamaa (Laurea-ammattikorkeakoulu / Theseus)
- Mari Myllyharju (Valtioneuvoston kanslia / Valto)
- Anita Niskanen (Valtion taloudellinen tutkimuskeskus / Doria)
- Ilkka Nokelainen (LUT-yliopisto / LUTPub)
- Erkki Nurmi (Taideyliopisto / Doria)
- Joonas Parviainen (Maanpuolustuskorkeakoulu / Doria)
- Mikko Pennanen (Turun yliopisto / UTUPub)
- Toni Raja-Hanhela (Lapin yliopisto / Lauda)
- Tiina Tolonen (Oulun ammattikorkeakoulu / Theseus)
- Jyrki Ilva, Kansalliskirjasto
- Reeta Kuuskoski, Kansalliskirjasto
- Erik Lindholm, Kansalliskirjasto
- Anis Moubarik, Kansalliskirjasto
- Emma Pietarila, Kansalliskirjasto
- Päivi Rosenström, Kansalliskirjasto
- Samu Viita, Kansalliskirjasto

Poissa

- Kaija Hirvi (Valtioneuvoston kanslia / Valto)
- Minna Liikala (Terveystieteiden tutkimuskeskus / Jukuri)
- Antti Nyqvist (Haaga-Helia ammattikorkeakoulu / Theseus)

Asialista

1. Kokouksen avaus
 - a. Pj. Merja Kallio avasi kokouksen
2. Asiantuntijaryhmän kokoonpano ja toimikausi
 - a. Merja Kallio (Vaasan yliopisto) aloitti ryhmän puheenjohtajana
 - b. Eetu Kejonen on tullut Tommy Lahtisen tilalle Åbo Akademin edustajaksi
 - c. Ryhmän kokoonpano ja toimikausi vahvistettiin [Tiedonhallinnan ohjausryhmän](#) kokouksessa 14.9.2020
 - i. Kokoonpanoa voidaan päivittää tarpeen mukaan
 - ii. Ryhmän toimikaudeksi vahvistettiin vuodet 2020-2023
3. Julkaisuarkistopalveluiden saavutettavuus / Reeta Kuuskoski
 - a. Suomenkieliset saavutettavuusselosteet saatiin kaikkiin arkistoihin 23.9.2020 mennessä eli työ saatiin valmiiksi lain asettaman määräajan puitteissa
 - i. Reeta kiitti kaikkia prosessissa mukana olleita hyvästä yhteistyöstä
 - b. Selosteiden ruotsin- ja englanninkieliset käännökset tehdään syksyn aikana
 - i. Kansalliskirjasto kääntää selostepohjan jolle tiedot laitetaan; aineistokohtaisten tietojen käännökset asiakasorganisaatioiden vastuulla
 - c. Myös paljon käyttöliittymien saavutettavuuspuutteita saatiin korjattua
 - i. Jäljellä olevien käyttöliittymäpuutteiden korjaus, testaus ja selosteiden päivittäminen jatkuu
 - d. Sisältöjen ja niihin liittyvien vastuiden osalta Celialta kysyttiin tarpeen mukaan lisätietoja
 - i. Vastaukset tulivat yleensä nopeasti.
 - e. Oppilaitosyhteistyötä on tehty opinnäytetöiden saavutettavuusasioissa / Tapani Lehtilä
 - i. Helsingin yliopisto oli kutsunut koolle kokouksen, jossa olivat mukana myös Jyväskylän yliopisto, MPKK, Oulun yliopisto, Tampereen yliopisto ja Turun yliopisto
 1. Kirjasto saa opinnäytteestä valmiin pdf:n johon on vaikea enää tehdä muutoksia, eli opinnäytteiden tekijät ja ohjaajat pitää saada mukaan prosessiin
 2. Koneellisia validaattoreita löytyy mutta ne vaativat täydellisyyttä, minkä vuoksi opinnäytteitä on käytännössä vaikea saada niistä läpi
 3. Kaivattiin tietoa siitä, mikä on opinnäytteiltä käytännössä vaadittava saatavuustaso
 - a. Työn pohjana [HY:n opiskelijoiden saavutettavuusohjeet](#), jotka lähetettiin Celian kommentoitavaksi
 - b. Celian vastauksen mukaan ohjeistus riittää minimitason saavuttamiseen jos opiskelijat toimivat sen mukaan
 - c. Organisaatioilla ei ole resursseja kattavaan jälkikäteistarkastukseen
 - d. PDF-UA-formaattia pidettiin opiskelijoiden kannalta liian vaativana
 - ii. Asiantuntijaryhmä jäi pohtimaan sitä riittääkö minimitaso ja tarvitaanko tarkistuksia
 1. Todettiin, että saavutettavuuden kanssa on vielä paljon tekemistä
4. Plan S:n vaatimukset julkaisuarkistoille / Mikko Pennanen
 - a. Eurooppalaisten tutkimusrahoittajien [Plan S -aloite](#) asettaa vaatimuksia julkaisuarkistoille

- b. Mikko alusti aiheesta Turun yliopistossa tehdyn selvitystyön pohjalta (ks. [esityskalvot](#))
- i. Plan S edellyttää, että kaikki mukana olevilta rahoittajilta rahoitusta saaneet uudet artikkelit ovat avoimia vuodesta 2021 lähtien
 1. Ei koske tässä vaiheessa kirjoja tai esim. tutkimusdataa
 2. Tavoitteena välitön avoimuus joko kultaista tai vihreää reittiä
 3. Hybridi-OA on sallittua vain transformatiivisen sopimuksen puitteissa vuoteen 2024 asti jatkuvan siirtymäkauden ajan
 - ii. Julkaisuarkistoihin asetettu laatuvaatimuksia, jotka Mikko jakoi kahteen ryhmään
 1. Makrotason vaatimukset
 - a. Julkaisuarkiston tulee olla rekisteröity OpenDOAR-palveluun
 - b. Julkaisuarkiston tulee tukea pysyviä tunnisteita (PID), kuten DOI, URN tai Handle. (PlanS suosii DOI:ta)
 - c. Julkaisuarkiston tulee olla vakaasti saatavilla (käytössä 99,7% ajasta)
 - d. Arkiston ylläpidon yhteystietojen tulee olla näkyvissä, yhteydenottoihin tulee vastata korkeintaan työpäivän viiveellä
 2. Mikrotason vaatimukset
 - a. Korkealuokkainen metadata saavutettavassa muodossa (tekninen saavutettavuus)
 - i. Tieto sekä toimitetun että alkuperäisen julkaisun PID:stä arkistossa olevassa julkaisussa, tulee olla myös merkittynä julkaisuun (dokumenttiin).
 - ii. Tieto OA-statuksesta arkistossa olevassa julkaisussa, tulee olla merkittynä myös julkaisussa (dokumentissa)
 - iii. Tieto OA-lisenssistä, tulee olla merkittynä myös julkaisuun (dokumenttiin)
 - iv. Metadatatassa tulee olla tieto cOAlition S-rahoittajasta (nimi ja rahoituspäätöksen tunnus)
 - b. CC0-lisensiointi tai public domain (Juridinen saavutettavuus)
 3. Lisäksi suositeltuja lisäkritereitä, jotka eivät pakollisia
 - a. Mm. aineistojen tallentaminen koneluettavassa formaatissa (esim. JATS/XML), kirjoittajien ja organisaatioiden tunnisteet, avoimet viittaustiedot
 - iii. Mikko oli koonnut pakollisista vaatimuksista taulukon, jonka toivoi olevan hyödyksi muillekin
 1. Makrotason vaatimuksiin ei liity suurempia ongelmia, sillä julkaisuarkistot täyttävät ne jo nyt
 2. Mikrotason vaatimuksiin liittyy sen sijaan hankalampia kysymyksiä
 - a. Tietojen tallentaminen pdf-muotoisena saatuaan dokumenttiin ei välttämättä käytännössä mahdollista, etenkin jos vaaditaan koneluettavuutta
 - i. Jyrki Ilva huomautti, että JATS/XML-formaatti oli alun perin mukana pakollisissa vaatimuksissa, eli vaatimus koneluettavuudesta saattaa olla jääne tästä
 - b. Julkaisuarkistojen metadatasuosituksessa on kyllä kentät tarvittaville tiedoille, mutta uusien tietojen kerääminen ja tallentaminen kattavasti (esim. rahoittajatieto) vaatii paljon lisätyötä ja prosessien kehittämistä
 - c. Rinnakkaistallenteiden osalta edellyttää monessa organisaatiossa työtä myös tutkimustietojärjestelmän päässä (ammattikorkeakouluilla Justuksessa), koska tallennus tapahtuu siellä
 - iv. Mikko kävi esityksessään läpi myös CC-lisenssien ja rightstatement-lausekkeiden käyttöä julkaisuarkistoissa
 1. Todettiin, että Plan S:n lisenssejä koskevat vaatimukset ristiriidassa esim. Elsevierin lisenssikäytäntöjen kanssa
 - c. Asiantuntijaryhmän keskustelussa todettiin, että aihepiiristä kaivattiin lisätietoja ja koulutusta
 - i. Todettiin, että Mikon esitys tulee verkkoon saataville
 - ii. Pohdittiin mahdollisuuksia järjestää syksyn aikana erillinen aiheita käsittelevä tilaisuus
 - iii. Jyrki lupasi tiedustella Suomen akatemiasta olisiko yhteistyössä heidän kanssaan mahdollista järjestää julkaisuarkistojen ja tutkimustietojärjestelmien parissa työskenteleville ihmisille suunnattu käytännönläheinen tilaisuus Plan S:n vaatimuksista
5. Muut asiat
- a. Annif otettu käyttöön useissa julkaisuarkistoissa / Samu Viita, Anis Moubarik
 - i. Demovideo Annifin käytöstä: <http://urn.fi/URN:NBN:fi-fe2020042219855>
 - ii. Osuvan, Trepon ja Theseuksen kokemukset olleet positiivisia
 - iii. Opiskelijat ovat osanneet toimia Annifin kanssa ja asiasanoitusta koskevat kysymykset ovat pikemminkin vähentyneet kuin lisääntyneet
 - iv. Theseuksen syöttölomakkeen teksteihin kaivattiin vielä viilausta
 - b. Kokotekstillisten tietojen merkkaukset ja niiden OAI-filtriointi
 - i. Jukurissa otettu käyttöön automaattinen kokotekstien merkkaukset
 - ii. Kokotekstiksi tulkitaan DSpace-tietue, jossa on vähintään yksi julkinen tiedosto (ORIGINAL-bundlessa)
 1. Jos mukana on pelkkiä abstrakteja sisältäviä tiedostoja, nekin tulkitaan kokoteksteiksi
 - iii. Jos kokoteksti löytyy, metadataan merkitään dc.format.bitstream=true, jos ei, arvo on false
 1. Esimerkkihaku käyttöliittymän puolella: https://jukuri.luke.fi/discover?rpp=10&etal=0&query=dc.format.bitstream:true&group_by=none&page=1
 2. OAI-filtri: <https://jukuri.luke.fi/oai/fulltext?verb=ListRecords&metadataPrefix=kk>
https://jukuri.luke.fi/oai/fulltext?verb=ListRecords&metadataPrefix=qdc_finna
 - iv. Rajausta pelkkiin kokoteksteihin voidaan jatkossa hyödyntää mm. Finna-haravoinnissa
 - c. Kansainvälisen DSpace-yhteisön kuulumisia / Jyrki Ilva
 - i. DSpace 7:n beta-versio 4 valmistumassa lokakuun alussa, etenemistä voi seurata täällä: <https://github.com/orgs/DSpace/projects>
 1. Beta-versio 3 ilmestyi heinäkuun alussa, beta 4:n kanssa kestänyt lomakauden takia pidempään
 2. Versio 7.0 julkaistaan beta 5:n valmistuttua, muutaman viikon testijakson jälkeen, todennäköisesti loppuvuodesta 2020 tai alkuvuodesta 2021
 3. Versioon 7.1 tulossa vielä lisää uusia ominaisuuksia
 - ii. Kansalliskirjasto ottaa version 7 käyttöön aikaisintaan vuonna 2022
 1. Ohjelmistossa tapahtuvien uudistusten takia vaatii etukäteen paljon selvitystä ja suunnittelua
 - d. Kansalliskirjaston asiakasorganisaatioille suunnattu [palvelukysely käynnissä](#)
 - i. Kyselyssä voi valita mitä Kansalliskirjaston palveluita koskeviin kysymyksiin vastaa
 - ii. Vastausaikaa 9.10. saakka
6. Seuraava kokous
- a. Seuraava kokous järjestetään torstaina 3.12. klo 9-11