

HUOM! Tämä on alkuperäisen artikkelin rinnakkaistallenne. Rinnakkaistallenne saattaa erota alkuperäisestä sivutukseltaan ja painoasultaan.

Käytä viittauksessa alkuperäistä lähdettä:

Tolonen, T., Nyqvist, A. & Marjamaa, M. (2022). Theseuksen julkaisutietojen siirtyminen ja metadatan avoimuus. Kreodi 4/2022. <http://urn.fi/URN:NBN:fi-fe2022120769761>

PLEASE NOTE! This is an electronic self-archived version of the original article. This reprint may differ from the original in pagination and typographic detail.

Please cite the original version:

Tolonen, T., Nyqvist, A. & Marjamaa, M. (2022). Theseuksen julkaisutietojen siirtyminen ja metadatan avoimuus. Kreodi 4/2022. <http://urn.fi/URN:NBN:fi-fe2022120769761>



Copyright: © 2022 by the authors and AMKIT-konsortio. Licensed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

Theseuksen julkaisutietojen siirtyminen ja metadatan avoimuus

07.12.2022



ÄKKISYVÄÄ, HIETANIEMEN RANTA.

Theseuksen käyttörajatun kokoelman viimeisin eli neljäs aalto käynnistyi elokuussa. Nyt kun ammattikorkeakoulut ovat tutustuneet tarkemmin uuteen toimintamalliin, on herännyt kysymyksiä metadatan avoimuudesta ja haravoinnista. Julkaisujen osalta replikoitumista Theseuksesta Finnaan on parannettu, opinnäytetöiden osalta replikointi säilyy ennallaan.

Theseus-toimisto selvittää koko ajan opiskelijoilta ja henkilökunnalta tulevia kysymyksiä. Osaan kysymyksistä löytyy nopeasti vastaus, esimerkiksi URNin toimimattomuus tai siirto kokoelmien välillä selviävät yleensä rutiinivastauksella. Välillä isommat ja laajemmat kysymykset vaativat kuukausien tai jopa vuosien selvittämistyön. Tässä artikkelissa tarkastellaan syksyn ajankohtaisia asioita.

Theseuksen julkaisujen siirtyminen Finnaan

Theseukseen tallennetut julkaisut ovat näkyneet Finnassa sellaisissa aineistotyypeissä, jotka eivät ole välttämättä kuvanneet aineiston tyyppiä kovinkaan hyvin. Rinnakkaistallennetut artikkelit ovat näkyneet kirjoina ja omat julkaisusarjat muina tutkimusjulkaisuina. Tämä on tuonut hakuihin erikoisia tuloksia tai jättänyt hyviä tuloksia pois. Nyt Theseukseen tallennettujen julkaisuiden näkymistä Finnan aineistotyyppirajaimessa on korjattu ja korjaukset otettu käyttöön Finnassa.

Korjaus tehtiin siten, että jatkossa Finnaan Theseuksesta haravoitujen julkaisujen aineistotyyppin määrittelemiseen käytetään hyväksi Theseuksessa olevaa metadatatietoa julkaisun OKM:n julkaisutyyppiluokituksesta. Tämä tarkoittaa sitä, että julkaisujen aineistotyyppit tulevat näkymään täsmällisemmin ja hienojakoisemmin kuin ennen. Kokonaisuudessaan Theseuksen julkaisujen aineistotyyppien tunnistus paranee uuden mappauksen myötä. Opinnäytteet käsitellään samalla periaatteella kuin ennenkin.

Mappaus Finnan aineistotyypeihin menee jatkossa käytännössä näin:

Jos haravoinnissa löytyy metadatatista kenttä `dc.type.okm`, sitä käytetään käytetään alla olevan taulukon mukaan.

Jos julkaisun metadatatassa ei ole `dc.type.okm`-kenttää, aineistotyyppin määrää tietueen ensimmäinen `type`-kenttä. Tämä on useimmiten `dc.type.ontasot` ja tämän kentän merkinnöille löytyy yleensä ennestään toimiva mappaus.

Ensimmäinen löytyvä kenttä voi olla myös jokin muu kuin `dc.type.ontasot` ja/tai sen sisältö voi olla sellainen, jolle ei ole valmista mappauksia. Jos Finna ei tunnista aineistotyyppiä, se mapataan Finnan aineistotyyppiin Teksti, muu > teksti. Tämä ei poikkea mitenkään vanhasta käytännöstä, vaan aina tulee satunnaisia julkaisuja, joiden tyyppiä ei tunnisteta.

Alla olevasta taulukosta (taulukko 1) näkyy, kuinka julkaisutiedonkeruun julkaisutyyppit Theseuksen `dc.type.okm`-kenttien sisällöstä on viety aineistotyypeiksi Finnassa.

Julkaisutiedonkeruun julkaisutyyppi	Aineistotyyppi Finnassa
A Vertaisarviodut artikkelit	
A1 Alkuperäisartikkeli tieteellisessä aikakauslehdessä	Artikkeli
A2 Katsausartikkeli tieteellisessä aikakauslehdessä	Artikkeli
A3 Kirjan tai muun kokoomateoksen osa	Kirjan osa
A4 Artikkeli konferenssijulkaisussa	Artikkeli
B Vertaisarvioimattomat tieteelliset kirjoitukset	
B2 Kirjan tai muun kokoomateoksen osa	Kirjan osa
B3 Vertaisarvioimaton artikkeli konferenssijulkaisussa	Artikkeli
C Tieteelliset kirjat (monografiat)	
C1 Kustannettu tieteellinen erillisteos	Kirja

Julkaisutiedonkeruun julkaisutyyppi	Aineistotyyppi Finnassa
C2 Toimitettu kirja, kokoomateos, konferenssijulkaisu tai lehden erikoisnumero	Kirja
D Ammattiyhteisölle suunnatut julkaisut	
D1 Artikkelit ammattilehdessä	Artikkeli
D2 Artikkelit ammatillisessa käsi- tai opaskirjassa, ammatillisessa tietojärjestelmässä tai oppikirja-aineisto	Kirjan osa
D3 Artikkelit ammatillisessa konferenssijulkaisussa	Artikkeli
D4 Julkaistu kehittämis- tai tutkimusraportti taikka -selvitys	Tutkimusraportti
D5 Oppikirja, ammatillinen käsi- tai opaskirja taikka sanakirja	Kirja
D6 Toimitettu ammatillinen teos	Kirja
E Suurelle yleisölle suunnatut julkaisut	
E1 Yleistajuinen artikkeli, sanomalehtiartikkeli	Artikkeli
E2 Yleistajuinen monografia	Kirja
E3 Toimitettu yleistajuinen teos	Kirja

Taulukko 1. Kuinka julkaisutiedonkeruun julkaisutyyppit vastaavat Finnan aineistotyyppijä.

Uusi toteutus parantaa selkeästi tarkkuutta muiden aineistojen kuin opinnäytetöiden osalta.

Theseuksen e-kirjat mappautuvat edelleen tällä hetkellä kirjoiksi eivätkä e-kirjoiksi, sama pätee pitkälle muihinkin julkaisuarkistojen e-kirjoihin. Mappausta testatessa Theseuksesta löytyi myös vähän yllättäen muutamia I1-tyypin julkaisuja. I-tyyppi voi sisältää niin monia erilaisia julkaisuja podcasteista verkko-oppimateriaaleihin, että näiden sähköisten julkaisujen vähäisyyden vuoksi tyydyttiin suhteellisen yleiseen aineistotyyppiin "Muu/määrittelemätön > Elektroninen".

Jos oman kirjaston Finna-näkymässä on muiden ammattikorkeakoulujen Theseus-kokoelmia ja niille on tehty Esisuodatus aineistotyyppin perusteella, kannattaa tarkistaa suodatusasetukset Finnan hallintaliittymässä. Viimeisin tieto muutoksesta on [Finna keskustelufoorumilla](#).

Myös pääsytietoja on määritelty nousemaan Finnaan Theseuksesta: restricted access ja open access -merkinnät nousevat nyt myös Finnaan. Embargotietojen nousemista Finna-toimisto tutkii edelleen.

PAS-siirrot

Kansalliskirjasto ja CSC ovat aloittaneet siirtoja CSC:n pitkäaikaispalveluun PASiin monografia-aineistolla. Testeissä on havaittu pieni määrä PDF- ja ePub-tiedostoja, jotka eivät ole menneet läpi CSC:n validoinnista salauksista tai vanhoista versioista johtuen. Testaamista jatketaan edelleen ja seuraavaksi testataan Mikkelin digitointiaineistojen lähetystä PASiin.

Theseuksen PAS-siirrot alkavat, kun järjestelmää on testattu muulla aineistolla riittävästi ja siirrot toimivat.

Palvelunestohyökkäysten torjunta

Korkeakoulujen tietoturvasuhteisuus on noussut tärkeäksi asiaksi tänä vuonna ja myös Theseuksen osalta on tullut kysymyksiä tietoturvasuhteisuudesta. Elokuussa YLE uutisoi Julkariin kohdistuneesta venäläisestä palvelunestohyökkäyksestä <https://yle.fi/a/3-12571164>. Myös Theseukseen on kohdistunut palvelunestohyökkäyksiä. Kansalliskirjasto seuraa tilannetta ja panostaa tietoturvaan ja selvittää tietoturvasuhteisuutta lisäävien palvelujen käyttöönottoa Theseuksenkin toiminnan taustalle.

Osana tätä tietoturvasuhteisuuden parantamisen kokonaisuutta Theseuksesta, kuten muistakin Kansalliskirjaston ylläpitämistä julkaisuarkistoista, on poistettu yhteydenottolomake. Sen kautta tuli Theseus-toimistoon paljon erilaista roskapostia ja kalasteluyrityksiä sekä silkkää Theseuksen palvelimia ja toimiston sähköpostilaatikoita kuormittavaa spämmiä. Pahimmillaan viestejä tuli toimistoon yhteydenottolomakkeen kautta muutaman minuutin sisällä useita satoja.

Opinnäytetöiden metadatan avoimuus

Käyttörajattu kokoelma on ammattikorkeakoulukentällä nostanut Theseuksen alkuaikojen avoimuuden peruskysymykset uudestaan esille: Millä perusteella opinnäytetyö on julkinen ja miksi opinnäytetyön metadatan pitäisi olla julkista? Voiko metadatan piilottaa? Theseuksesta metadata replikoituu niin Finnoihin kuin tulee haravoiduksi avoimen OAI-PMH-rajapinnan kautta Coreen ja OpenAireen.

Perustuslaki, julkisuuslaki, Pohjois-Suomen hallinto-oikeuden päätös vuodelta 2019 sekä Opetus- ja kulttuuriministeriön kirje ammattikorkeakouluille 2004 ovat lähtökohtia Theseuksen avoimuudelle. Viime kädessä ammattikorkeakoulut ovat itsenäisiä oikeustoimihenkilöitä, jotka tekevät itse päätökset omista toimintamalleistaan ja prosesseistaan. Niiden täytyy siis huolehtia, että avoimuuden periaatteet ovat selkeitä myös opinnäytetöitä ohjaaville opettajille.

Koko opinnäytetyön avaaminen kokonaisuudessaan vapaaseen internetiin edellyttää opiskelijalta saatavaa suostumusta tekijänoikeuslain nojalla; tätä Theseuksessa pyydetään opiskelijalta tallennuksen yhteydessä.

AMKIT-konsortio on tulkinnut metadatan avoimuuden pohjautuvan perustuslain vaatimukseen siitä, että kenellä tahansa tulee olla oikeus saada tieto julkisesta asiakirjasta. Sen sijaan tietosuojavaltuutetun kannanoton perusteella julkisuuslain pohjalta on mahdollista tiedottaa oppilaitoksesta valmistuneista opinnäytetöistä niiden perustietoja koskevin osin ilman erityisen suostumuksen hankkimista ([julkisuuslaki 16.3 §](#), [Opinnäytetyöt ja tietosuoja](#) 2010: 5).

Avoin tiede ja tutkimus -yhteistyössä käynnistyi lokakuussa Opinnäytetöiden avoimuuden

linjauksen luonnostelutyöryhmä, jossa luodaan suositus opinnäytetöiden avoimuudelle. Työryhmässä on edustus niin AMKIT-konsortioista kuin Theseus-toimistosta. Toivomme, että linjauksessa otetaan kantaa niin avointen lisenssien suositukseen kuin metadatan avoimuuteenkin. Odotamme linjauksen valmistumista ensi vuoden lopulla suurin toivein.



Kuva 2. Theseus-toimisto vietti jouluillallista Bistro Bardot'ssa. Tavattiin ensimmäistä kertaa livenä moneen vuoteen. Kuva: Tiina Tolonen.

Vuoden 2023 suunnitelmia

Kansalliskirjasto käynnistää tulevana vuonna sen ylläpitämässä julkaisuarkistoissa DSpace-versiopäivityksen. DSpace 7.4. on valmistunut lokakuussa 2022, versio 7.5 ilmestyy helmikuussa 2023, 7.6. on vuorossa seuraavana kesänä. Versiopäivitys sisältää suuria muutoksia käyttöliittymään ja tulossa on myös parannuksia käyttäjähallintaan. Päivitys aloitetaan ensin pienemmistä arkistoista ja viimeisenä päivitysvuorossa on Theseus, kunhan mahdolliset päivityksen aiheuttamat ongelmakohdat on havaittu ja korjattu. Aikataulullisesti Theseuksen päivitys menee todennäköisimmin vuodelle 2024.

Vuonna 2023 Theseus-toimistossa tehdään markkinointimateriaaleja Theseusta varten ja seurataan DSpace-päivityksen etenemistä ja PASin käyttöönottoa. Roll upin teosta ammattikorkeakoulujen käyttöön on alustavasti sovittu graafikon kanssa. Theseus-toimisto reagoi tekijänoikeuslain käyttöönottoon sekä seuraa saavutettavuusdirektiivin etenemistä ja tiedottaa näistä eteenpäin. Käyttörajatun kokoelman käyttäjätyytyväisyydestä tehdään kysely keväällä 2023.

Konsortion työryhmät esittäytyvät tammikuussa

Tammikuun loppupuolella 26.1.2023 AMKIT-konsortio järjestää Konsortion työryhmät tutuksi - tapahtuman verkossa.

Tilaisuudessa esittäytyvät:

- Kestävä kehitys -ryhmä
- Kreodin toimitus
- Metatietoryhmä
- Pedagogiikkaryhmä ja Avoimet oppimateriaalit -ryhmä
- Theseus-toimisto.

Tervetuloa mukaan päivään!

Kirjoittajat



Tiina Tolonen

Informaatikko

Tiina Tolonen on informaatikko Tiedekirjasto Pegasuksessa, joka palvelee Oulun yliopistoa ja Oulun ammattikorkeakoulua.

[Kirjoittajan muut artikkelit](#) >



Antti Nyqvist

tietoasiantuntija

Haaga-Helia

[Kirjoittajan muut artikkelit](#) >



Minna Marjamaa

tietoasiantuntija

Laurea-ammattikorkeakoulu

[Kirjoittajan muut artikkelit](#) >

Artikkelin tiedot

Kirjoittaja: Tiina Tolonen, Antti Nyqvist, Minna Marjamaa

Numero: 4/2022

Kuvat: CC BY 4.0

URN: <http://urn.fi/URN:NBN:fi-fe2022120769761>

Lisenssit



Tämä teos on lisensoitu [Creative Commons Nimeä 4.0 Kansainvälinen -lisenssillä](#).