

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 yksityiskohtien kevät. Kreodi 2/2022. <http://urn.fi/URN:NBN:fi-fe2022060844437>

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 yksityiskohtien kevät. Kreodi 2/2022. <http://urn.fi/URN:NBN:fi-fe2022060844437>



**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 yksityiskohtien kevät

08.06.2022

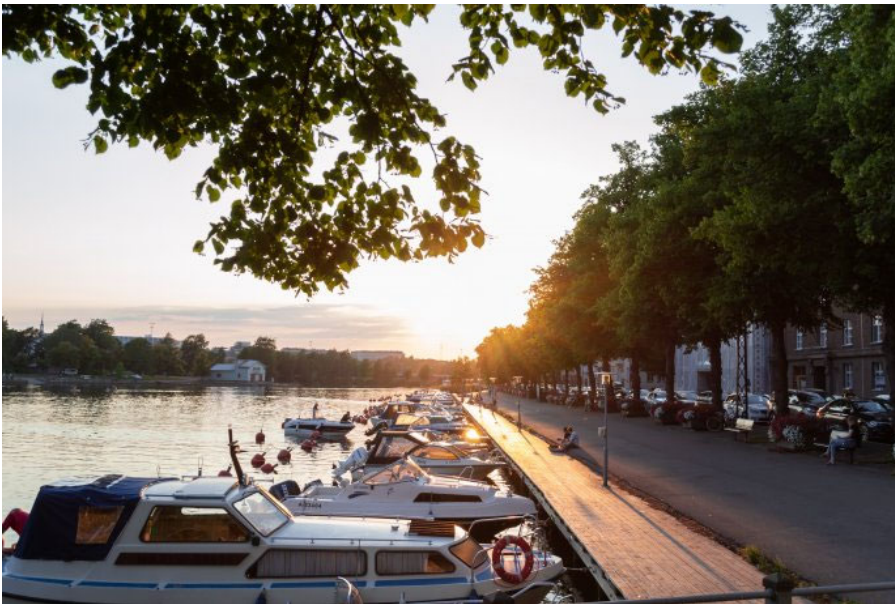


PHOTO BY [ABBILYN ZAVGORODNIAIA](#) ON [UNSPLASH](#)

Theseus-työhön on tänä keväänä sisältynyt runsaasti yksityiskohtien selvittämistä. Käyttörajattuun kokoelmaan liittyviä käytänteitä on ratkottu ja metadataan massaeditoitu tieto käyttörajauksesta. Theseus-Finna-ongelmia on selvitetty ja ratkaisujakin on löydetty. Myös periaatteellista keskustelua on käyty.

## Käyttörajatusta kokoelmasta tieto metadataan

Amk-kirjastoista on kyselty, kuinka voisi saada tilastotietoa käyttörajatun kokoelman tallennuksista kuukausittain. Osin tätä tarkoitusta varten ja myös rajattujen tallennusten helpompaa erottamista varten tänä keväänä on lisätty Theseuksen metadataan

uusi kenttä dc.rights.accessrights ja sille arvoksi Rajattu käyttöoikeus / Restricted access / Tillgången är begränsad.

Uusi kenttä lisättiin item templateihin, josta kentän arvo tulee uusiin käyttörajattuihin tallennuksiin automaattisesti. Lisäksi se massaeditoitiin tuota aiemmin käyttörajattuihin kokoelmiin tallennettuihin töihin, joten se löytyy nyt kaikista käyttörajatuista tallennuksista. Jos kirjastoissa tehdään siirtoja kokoelmasta toiseen, täytyy tämä kenttä joko lisätä tallennukseen tai poistaa siitä. Tästä löytyy ohje [Theseuksen eDuuni wikistä](#).

Kuukausittaisia tallennusmääriä käyttörajatuista opinnäytetöistä voidaan hakea Theseuksesta seuraavalla hakulausekkeella, esimerkissä haetaan helmikuussa 2022 käyttörajattuun kokoelmaan tallennettuja töitä: dc.date.accessioned:[20220201 TO 20220228] Valitkaa ensin käyttörajattu kokoelma, ja hakulaatikon alta ”tämä kokoelma”. Haussa ei saa olla kaksoispisteen jälkeen välilyöntejä ja haluttu kuukausi ja päivät merkitään aina kahdella numerolla.

Theseus-kehittämissryhmässä (muistiot eDuunissa) on myös keskusteltu, olisiko järkevää edistää tiedon viemistä opinnäytetöiden avoimuudesta tai käyttörajauksesta myös kansalliseen tilastointiin Vipusen kautta. Vipusesta tällä hetkellä löytyy opinnäytetöiden osalta vain tieto siitä, ovatko opinnäytetyöt hankkeistettuja vai eivät. Toiseksi tätä ei viedä eteenpäin, syynä se, että Avoimen tieteen ja tutkimuksen seurannassa jäi pois lopulta kysymys opinnäytetöiden avoimuuden asteesta. Seuraamme tilanteen kehitystä.

## Theseus-Finna -kysymyksiä

Tänä keväänä AMKIT-konsortio ja Finna-toimisto ovat järjestäneet useamman yhteisen Finna-aiheisen palaverin. Kokouksissa on käyty läpi amk-kirjastokentältä kerättyjä Finnaan liittyviä ongelmakohtia. Theseuksen osalta käsiteltävänä on ollut kolme pääkohtaa.

### Theseus-Finna-ongelmakohdat:

- Kaikki dc.subject-alkuiset kentät siirtyvät Finnaan asiasanoina. Erilaiset dc.subject-alkuiset kentät sisältävät tietoa opinnäytteen koulutusaloista, suuntautumisesta ja asiasanoista. Jatkossa pyritään siirtämään Theseuksesta Finnaan vain dc.subject.yso ja dc.subject.degreeprogram-kentät ja nämä molemmat oman otsikkonsa alle. Dc.subject.discipline ja dc.subject.specialization

-kentät pyritään jättämään kokonaan pois Finnan tietuenäkymästä. Finna-toimisto selvittää edelleen tämän käytännön toteutusta.

- **Käyttörajatun kokoelman töiden näyttäminen Finnassa.** Tietueet, joissa on jossakin dc.rights-kentässä arvo Rajattu käyttöoikeus / Restricted access / Tillgången är begränsad tunnistetaan nyt käyttörajatuksi. Tämä tarkoittaa, että ne eivät enää tule mukaan Finna.fi:n vapaasti verkossa saatavilla oleviin hakutuloksiin.

Samaa rajausta hyödyntää myös Finnan Oikopolku tutkittuun tietoon -palvelu, jonka kautta on tarkoitus löytää ainoastaan vapaasti verkossa olevat dokumentit.

- **Theseuksesta Finnaan haravoidut ammattikorkeakoulujen rinnakkaistallenteet ovat näkyneet Finnoissa kirjoina.** Jatkossa Theseukseen tallennettujen julkaisujen metadatatista, kentästä dc.type.okm, löytyviä Virta-tietomallin julkaisutyyppisiä hyödynnetään Theseuksesta haravoitujen julkaisujen luokittelemisessa eri aineistotyyppisiin Finnassa. Julkaisutyyppitiedon perusteella rinnakkaistallenteet näkyvät artikkeleina, kirjoina, kirjan osina, tutkimusraportteina tai väitöskirjoina faseteissa. Theseukseen on myös rinnakkaistallennettu taiteellisen ja taideteollisen toiminnan julkaisuja (julkaisutyyppi F) sekä AV-julkaisuja (julkaisutyyppi I1) joiden aineistotyyppi Finnassa on vielä määrittelemättä. Näiden ratkaisujen pitäisi toimia myös Virran tulevassa attribuuttipohjaisessa julkaisutyyppiluokittelussa.
- **Theseuksesta siirtyä Finnoihin useampi osoite,** mikä on ongelmallista erityisesti käyttörajatuissa töissä, sillä käyttäjä saattaa törmätä niiden IP-rajaukseen. Theseuksen paranneltu kirjautumissivu informoi käyttäjää aiempaa paremmin siitä, että hän on navigoinut käyttörajattuun tiedostoon. Myös dc.rights-kentän käyttöönotto parantaa tilannetta, sillä sen avulla käyttörajatun kokoelman työt saadaan rajattua Finnassa hakutuloksesta pois käyttämällä Verkossa saatavilla -rajausta.

Theseus-Finna-korjausten edistymistä seurataan syksyllä seuraavassa AMKIT-konsortion ja Finna-toimiston kokouksessa.

# Tiedostojen nimeäminen ja kuvailu

Theseuksessa esiintyy paikoin outoja tiedoston nimiä. Tämä johtuu siitä, että opiskelija on tallentanut tiedostonsa ohjeistuksen vastaisella nimellä tai lisännyt vapaaehtoiseksi määritellyn kuvailutiedon, joka näkyy tietueen etusivulla Theseuksessa tiedostonimen paikalla. Joskus kuvailuksi on tallennettu jopa koko opinnäytetyön tiivistelmä.

Erikoiset ja ohjeistuksen vastaiset nimeämiset aiheuttavat ongelmia Finna-haravoinnissa, sillä Kansalliskirjaston mukaan enkoodausta on vaikea tehdä jos on varauduttava kaikkiin mahdollisiin syötteisiin. Kansalliskirjasto toivoo tämän vuoksi, että jatkossa opiskelijoiden nimeämät tiedostot muokataan ohjeistetun mukaiseksi. Töiden nimiä ei kuitenkaan tarvitse muokata taannehtivasti.

## Ohjeistus on seuraava:

- Nimeä PDF/A-tiedostosi muodossa sukunimi\_etunimi, esimerkiksi "Opiskelija\_Otto".
- Jos tekijöitä on useampi, nimeä työ muodossa sukunimi\_sukunimi\_sukunimi.
- Sallitut merkit ovat 1-9, a-z, A-Z, \_, - ja välilyönti.
- Älä käytä tiedoston nimessä ääkkösiä.

Ammattikorkeakouluista on esitetty kysymys, olisiko tiedoston kuvailukenttä mahdollista poistaa tallennuslomakkeelta kokonaan. Se ei kuitenkaan ole mahdollista, sillä se on Kansalliskirjaston ylläpitämien julkaisuarkistojen yhteinen piirre ja tällä hetkellä koodi on jäädytetty seuraavaan DSpaceen versioon siirtymisen vuoksi.

# Syöttölomakkeen päivitykset kesällä

Keräämme parhaillaan kokoon tarvittavia syöttölomakkeen päivityksiä. Seuraavia asioita ollaan päivittämässä:

- ohje tiivistelmän tallentamiseen
- ohje tiedoston kuvailuun
- virhekorjauksia.

Nämä päivitykset on tarkoitus toteuttaa samalla alustapäivityksen kanssa. Jos toivot muutoksia syöttölomakkeelle, ole yhteydessä Theseus-toimistoon.

# Erilaiset tiedostomuodot

Voiko Theseukseen tallentaa muuta tiedostotyyppiä kuin PDF- ja PDF/a-tiedostoa? MSWord-tiedostojen (pääte on .doc ja .docx) tallentaminen Theseukseen on estetty, mutta sen sijaan ePub-tiedostoja voi kyllä tallentaa. Epub-tiedostomuoto on käytettävämpi ja sen saavutettavuusominaisuudet ovat pdf-tiedostoja parempia, joten se on julkaisuille hyvä vaihtoehto.

Tänä keväänä on myös keskusteltu äänikirjojen ja -artikkeleiden tallentamisesta Theseukseen. Miten tällainen teknisesti olisi Theseuksessa mahdollista julkaista? Mp3-tiedostot toimivat Theseuksessa, samoin muut [HTML5:n tukemat ääniformaatit kuten Mp4 ja WAV](#). Sen sijaan Celian tuotantovaatimusten mukaista [Daisy-äänitiedostoa](#) Theseukseen ei voi tallentaa.

## URN-generointi ja Justus

Vuoden ensimmäisillä amk-rinnakkaistallentajien aamukahveilla Emma Pietarila Kansalliskirjastosta kertoi URN-tunnuksista ja niiden käytöstä Theseuksessa. Keskustelu liikkui URNien haravoinnin toteutuksen arkkitehtuurin kuvauksesta yleisempiin pohdintoihin pysyvien tunnusten käytöstä.

Myös konkreettisia esimerkkejä käytiin läpi. Monessa ammattikorkeakoulussa on havaittu niin sanottu kahden URNin ongelma: Kun julkaisu, jolla on jo aiemmin URN-tunnus, rinnakkaistallennetaan Theseukseen, se saa rinnakkaistallennuksessa toisen URN-tunnuksen. Pysyvien tunnisteiden käytön idean kannalta tämä ei ole välttämättä hyvä tilanne, sillä samaa julkaisuja ei ole tarpeen identifoida kahteen kertaan eri tunnuksella. Toisaalta tämä on Theseukseen tallentaville amkeille selkein toimintatapa, ja vaihtoehtoiset toimintatavat tuovat mukanaan erilaisia epäkäytännöllisyyksiä.

Saman URN-tunnuksen käyttö sekä rinnakkaistallenteelle että alkuperäiselle julkaisulle aiheuttaa sen, että URN-tunnukseen pohjaavan linkin laskeutumissivu on se, minkä URN-harvesteri on viimeisimpänä poiminut. Kansalliskirjasto on tehnyt työtä, jotta pystyttäisiin paremmin tukemaan samalle URN-tunnukselle useita laskeutumissivuja.

Toisilla, toukokuussa järjestetyillä aamukahveilla liikuttiinkin jo

kauempana Theseuksesta ja keskusteltiin Justuksen ja julkaisutiedonkeruun asioista, erityisesti Justuksen tekijätietominaisuuden käytöstä ja käytännöistä. Vaikutelmaksi jäi, ettei kovinkaan moni ammattikorkeakoulu käytä Justuksen tekijätietominaisuutta. Niitä jotka sitä käyttävät, mietityttää tekijätietojen päivittämiseen ja niiden automatisointiin liittyvät käytännöt, erityisesti sitten kun tekijätietojen tietomalli Justuksessa laajenee tukemaan tietojen siirtämistä Tutkimustietovarantoon. Oletettavaa on, että kunhan tiedejatutkimus.fi-portaalin Tutkijan profiili-osio tulee mahdollisesti syksyllä käyttöön, myös kiinnostus Justuksen tekijätietominaisuuden käyttöön kasvaa.

## Kysymykset Theseus-toimiston sähköpostiin

Suuri osa toimistolle tulevista viesteistä tulee eri ammattikorkeakoulujen opiskelijoilta, jotka kyselevät joko tallennuksensa käsittelyaikaa tai haluavat tehdä korjauksia juuri tekemäänsä tallennukseen. Kevään aikana viestejä on tullut myös opiskelijoilta, jotka eivät ole pystyneet tekemään tallennusta siitä syystä, ettei heille ole tullut tallennuskokoelmaa valittavaksi Haka-kirjautumisen jälkeen. Olisi toivottavaa, että ammattikorkeakoulujen ohjeissa kerrotaisiin töiden julkaisun ja tarvittavien korjausten tai muutosten, kuten tiedoston vaihto, tapahtuvan nimenomaisesti opiskelijan oman ammattikorkeakoulun kirjaston kautta. Näin opiskelijat osaisivat ottaa yhteyttä suoraan oikeaan tahoon, sillä usealla yhteydenottajalla vaikuttaa olevan aikamoinen kiire tai hätä valmistumisensa suhteen.

Tallennusongelmissa suurin syy on ollut se, että opiskelija on avoimen amkin opiskelija eikä tutkinto-opiskelija. Oikeus tallentamiseen määräytyy opiskelijan affiliaation mukaan, joka tutkinto-opiskelijalla on "student" ja avoimen amkin opiskelijalla "affiliate". Tuolla tasolla ei ole riittäviä oikeuksia tallentamiseen. Jos tallentaminen ei onnistu, opiskelijan statuksen voi helposti testata näin:

Tarkista Hakan attribute-test palvelusta: <https://firmitas.csc.fi/haka/> että EduPersonScopedAffiliation -attribuutti on muotoa student@amk.fi (jossa amk on kyseisen amk:n tunnus eli esimerkiksi Oamkin tutkinto-opiskelijan muoto on (student@oamk.fi). Opiskelijan tulee näissä tapauksissa olla yhteyksissä opintotoimistoon, Theseus-toimisto ei pysty muuttamaan opiskelijan affiliaatiota.

# Theseuksen alustapäivitys kesällä

Theseuksen alustapäivitys tehdään Kansalliskirjaston ilmoituksen mukaan **20–21.6.2022**. Alustapäivitys näkyy asiakkaille siten, että 20.6. klo 12 alkaen kirjautuminen estetään Theseukseen, mutta muutoin Theseus on käytettävissä haun ja selauksen osalta. Kirjautuminen palautuu viimeistään 21.6. klo 14, jolloin Theseus on saatu siirrettyä uudelle alustalle.

## Käyttörajatun kokoelman kolmas aalto

Käyttörajatun kokoelman käyttöönoton kolmas aalto käynnistyy syyskuussa. Mukaan lähtevät VAMK, Humak, Centria, Novia ja Polamk. Tiistaina **6.9.2022** viikoittaisen huoltokatkon yhteydessä lisätään uudet ammattikorkeakoulut mukaan.

Kun kokoelmat on otettu käyttöön, uusien ammattikorkeakoulujen kannattaa testata kokoelmaa niistä ip-osoitteista jotka on ilmoitettu rajaukseksi ja myös niiden ulkopuolelta. Näin mahdolliset virhetilanteet havaitaan heti.

## Syksy

Syksyllä järjestetään käyttörajatun kokoelman kolmannelle aallolle yhteinen tilaisuus heti käyttöönoton jälkeen. Lisäksi järjestämme kevään tapaan Theseus-tuokion, jossa käsitellään ajankohtaisia asioita keskustellen.

---

## 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: 2/2022 Ohjauksen ja oppimisen teemanumero

URN: <http://urn.fi/URN:NBN:fi-fe2022060844437>

## Lisenssit



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