



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Please cite the original version: Tolonen, T. ; Marjamaa, M. & Nyqvist, A. (2019) Odotettu Justus–Theseus - integraatio viimeinkin avattu. *Kreodi* 3.

URL: <https://www.kreodi.fi/en/32/Artikkelit/565/Odotettu-Justus%E2%80%93Theseus--integraatio-viimeinkin-avattu.htm>

Odotettu Justus–Theseus -integraatio viimeinkin avattu

Tiina Tolonen
Minna Marjamaa
Antti Nyqvist

Kauan odotettu Justus –Theseus -integraatio avattiin maanantaina 6.5.2019. Integraatio yhdistää OKM:n julkaisutiedonkeruun ja rinnakkaistallentamisen: Jatkossa ammattikorkeakoulujen julkaisuarkisto Theseukseen voi syöttää julkaisuja suoraan julkaisutietojen tallennuspalvelusta Justuksesta. Tämä tarkoittaa Theseuksen kytkeytymistä tiiviimmin kansalliseen julkaisutieto-infrastruktuuriin ja avoimen tieteen prosessien virtaviivaistumista ammattikorkeakouluissa. Palvelu lähti toimimaan ja yhteyteen on toistaiseksi oltu melko tyytyväisiä. Tässä koostetta kevään muutostyöstä ja kysymyksiä herättäneistä asioista.

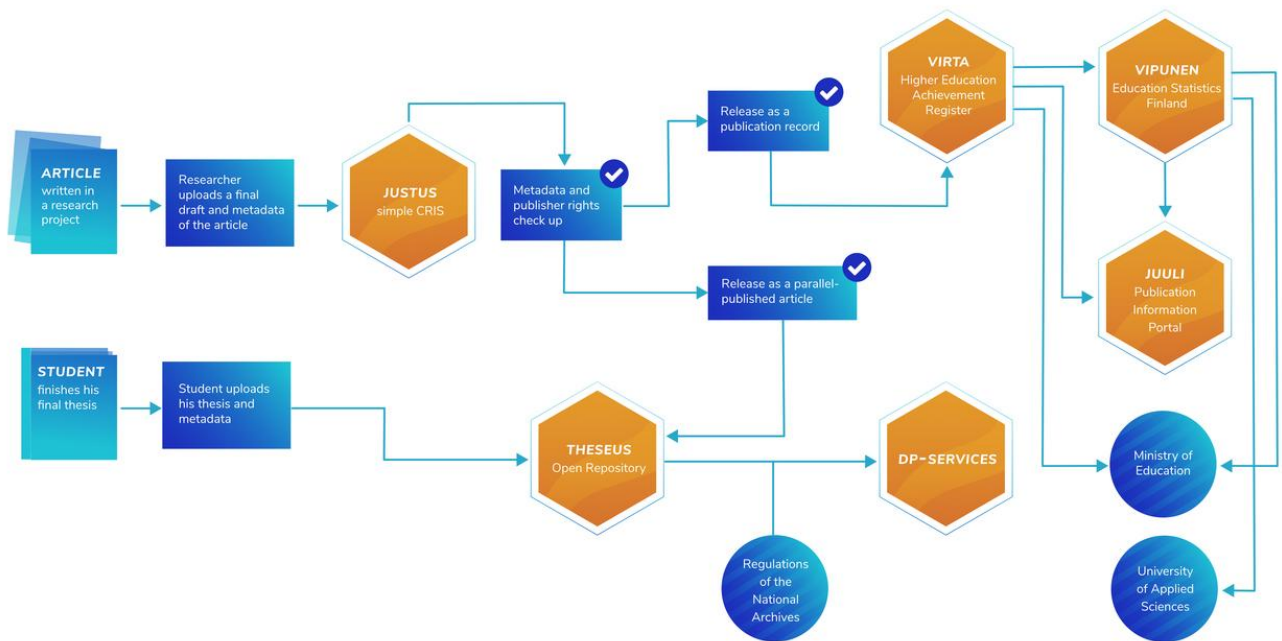
Tänä keväänä toteutettu neljä pitkään valmisteltua Theseus-uudistusta:

1. Metadatatäpäivitys: metadata uudistettu kansallisen metadatasuosituksen mukaiseksi
2. Kokoelmauudistus: jokaiselle ammattikorkeakoululle jäi ainoastaan yksi opinnäytetyö- ja julkaisukokoelma
3. Syöttömakeuudistus: Uusi Dspace-pohjainen syöttömake sekä opinnäytetöille että julkaisuille
4. Justus–Theseus -yhteys: Justukseen liittyneiden amkien osalta julkaisujen tallentaminen siirtyi tehtäväksi Justuksen kautta

Teknisellä puolella tapahtui parannuksia esimerkiksi palvelintekniikan osalta.

Mitä tästä kaikesta seurasi?

Theseuksen jo tähänkin saakka hyvin yhtenäinen metadata on toivottavasti vieläkin yhtenäisempää, tosin julkaisujen osalta saattaa vielä editointia olla tiedossa joidenkin kenttien osalta. Uudet Dublin Core -kentät mahdollistivat erilaisten selausfasettien luomisen. Nämä eivät vielä valitettavasti kuitenkaan näyttäydy kaikessa hienoudessaan, koska kokoelmien takautuva massaeditointi on vielä kesken. YAMK-töitä on nyt mahdollista selata omana kokonaisuutenaan, tästä on tullut toiveita vuosikausien ajan sekä amkeista että Theseuksen käyttäjiltä. Theseuksesta hakuja tehneet ovat saattaneet huomata vasteaikojen lyhentyneen, kun jatkuvasti kasvavaa käyttökuormaa on jakamassa useampi palvelin.



Kaaviokuva Theseuksen pääprosesseista: 1) julkaisun ja 2) opinnäytetyön tallennuksesta ja niiden kytkeytymisestä Virtaan ja PASiin. © Petra Boije

Kaksi kärpästä yhdellä iskulla: Rinnakkaistallennus Theseukseen Justuksen kautta

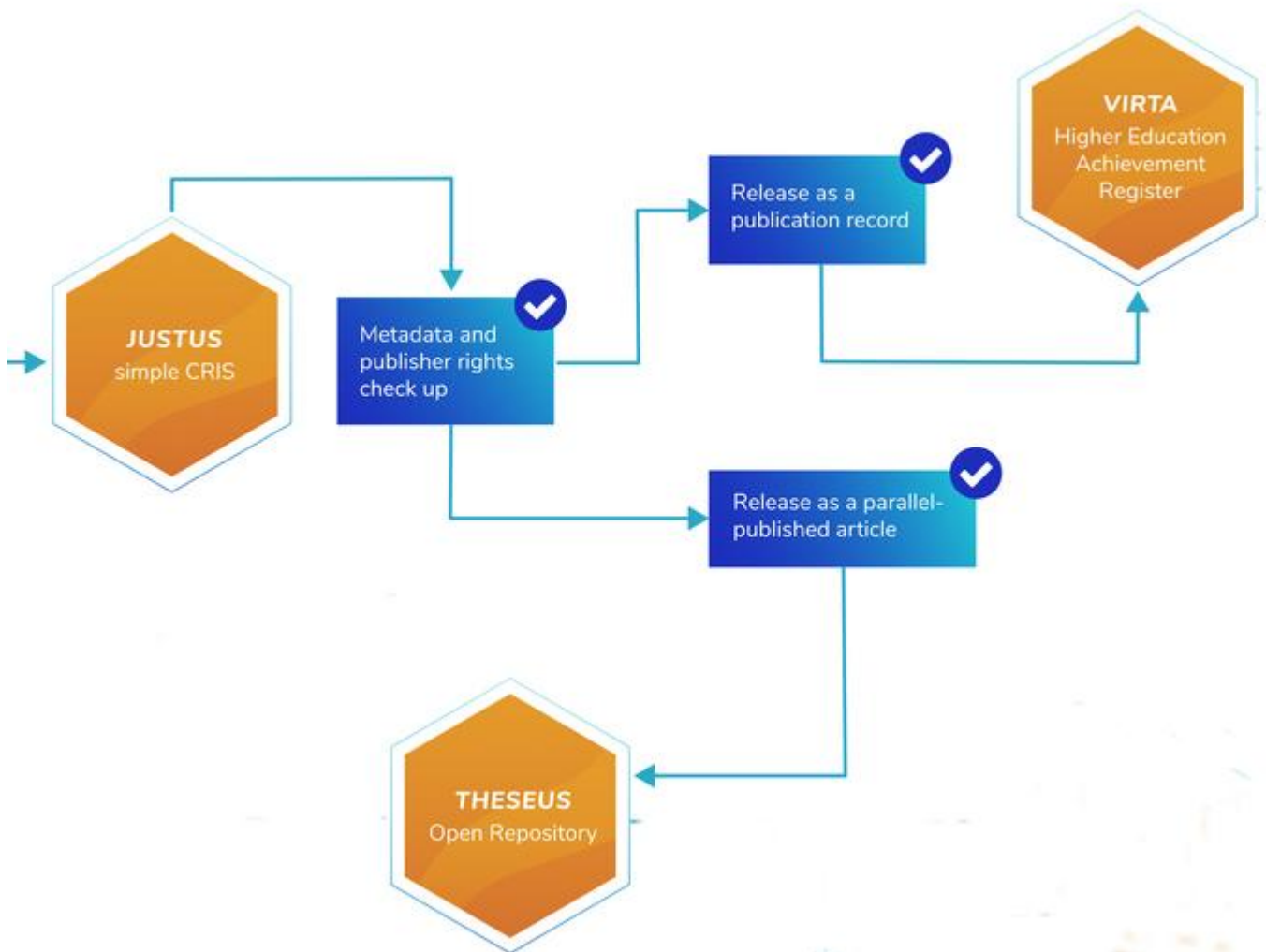
Suomen yliopistojen edellytetään julkaisevan vuoteen 2020 mennessä tiedejulkaisut 100 % avoimesti. Julkaisujen avoimuuden aste vaikuttaa myös rahoitukseen, sillä OKM:n ammattikorkeakoulujen uudessa rahoitusmallissa (2021–2024) on lisäkerroin 1,2 avoimesti saatavilla oleville julkaisuille (julkaisuluokat A–E).

Painetta siis julkaisujen avaamisen siis on. Ammattikorkeakoulujen väliset erot julkaisujen avoimuuden asteessa ovat isoja: osalla A-tyypin julkaisuista hyvin pieni osa avointa. Erityisesti A-tyypin julkaisujen avaamiseen auttaa rinnakkaistallentaminen, sillä vertaisarvioidut artikkelit julkaistaan edelleen isoksi osaksi suljetuissa lehdissä. Toisaalta rinnakkaistallentaminen kokoaa oman korkeakoulun julkaisut yhteen ja lisää niiden löydettävyyttä ja sillä tavoin viittausmääriä.

Rinnakkaistallennus on nyt tehtävissä samalla kuin OKM:n tiedonkeruu eli työmäärä vähenee. Justuksessa mukana oleville organisaatioille suositellaan julkaisujen tallentamista Theseukseen Justuksen kautta. Tämä nopeuttaa prosesseja ja tuo mukanaan tarpeellista metadattaa. Tieto kulkee Justuksesta Theseukseen, korjaukset tehdään Justuksen päässä niin Virtaan kuin Theseukseen.

Rinnakkaistallennusvastuu on tuonut Virta-yhdyshenkilöille eteen uusia kysymyksiä. On tärkeää osata perusasiat rinnakkaistallentamisesta ja ymmärtää, mitä on tekemässä, ja on myös kyettävä vastaamaan julkaisuilmoituksia tekevän henkilöstön kysymyksiin rinnakkaistallentamisesta. Tämä kannattaa huomioida omassa korkeakoulussa. Julkaisuilmoitusten tekeminen on harvemmin julkaisuille usein uusi ja monimutkaiseltakin tuntuva asia erilaisine julkaisuluokkineen ja muine

bibliografisine tietoineen. Rinnakkaistallenteiden lisenssi- ja embargokysymykset varmastikin nostavat yhteydenottojen määrää.



Rinnakkaistallentaminen jää Justus-pääkäyttäjän vastuulle. © Petra Boije

Korkeakoulun oman julkaisusarjan tallentaminen

Ensisijaisesti julkaisujen tallennus tehdään Theseukseen Justuksen kautta. Theseukseen suoraan tallentaminen on vain niille töille, joita ei haluta Virtaan. Miten toimitaan korkeakoulun omien julkaisusarjan kokoomateosten ja monografioiden kanssa? Nyrkkisääntönä on, että ne työt, jotka halutaan Virtaan ja soveltuvat tiedonkeruun kategorioihin, tallennetaan Justuksen kautta. Justukseen on mahdollisesti jatkossa tulossa myös kaatoluokkatallennus, jonka kautta voi tallentaa muitakin, esim. konferenssiesityksen ppt-tiedostoja.

Jos haluaa tallentaa suoraan Theseukseen, julkaisun tallentamista varten on ensin rekisteröidyttävä Theseukseen. Rekisteröityminen vaaditaan, koska toistaiseksi ei ole mahdollista Haka-kirjautumisen perusteella päätellä, keillä on oikeus tallentaa julkaisuja.

Miten saada oma julkaisusarja esille Theseuksessa?

Theseuksessa voi tallentaa julkaisuja ainoastaan yhteen kokoelmaan eli **Julkaisut**. Oman julkaisusarjan saa esille merkitsemällä oman julkaisusarjan nimen Justuksessa Syötä julkaisu - välilehdellä kenttään Julkaisusarja. Julkaisusarja on vapaasanakenttä, jossa voidaan määrittellä julkaisulle esimerkiksi organisaatiokohtainen julkaisusarja.

Tieto välittyy julkaisuarkistoon ja sitä käytetään julkaisusarjojen näyttämiseen. Tämä siirtyy Theseuksessa kenttään dc.type.other. Jos tallentaa suoraan Theseukseen, julkaisusarjan kentän voi merkitä suoraan kenttään dc.type.other.



The screenshot shows a form with three rows. The first row is labeled 'Oikeudet' and has a dropdown menu set to 'CC BY-SA 4.0 Nimeä - Jaa samoin' and a 'Tyhjennä' button. The second row is labeled 'Versio' and has a dropdown menu set to '-- Valitse --'. The third row is labeled 'Julkaisusarja' and has an empty text input field.

Julkaisusarjan nimen tallennus Justuksessa.

Huomioida kannattaa myös se, että jos tekee rinnakkaistallennuksia sekä Theseukseen suoraan että Justuksen kautta, ja haluaa niiden näkyvän samassa Julkaisulaijissa on Justukseen Julkaisusarja -kenttään syötettävän tiedon vastattava Theseuksen syöttölomakkeen kautta määrittyvää dc.type-other -tietoa, esim. XXX ammattikorkeakoulu Julkaisut tai fi=Rinnakkaistallennetut artikkelit|sv=Parallellpublicationer|en=Self-archived articles|.

PAS-siirrot ja opinnäytetöiden arkistointi, PDF/A

Opinnäytetöiden arkistoinnista on sovittu Arkistolaitoksen kanssa syksyllä 2017: Theseus-tallennus ja jatkossa valmistuva PAS-siirto riittää Theseuksessa olevien töiden osalta täyttämään vaatimukset sähköisestä arkistoinnista.

Miltä PAS-aikataulu näyttää nyt? Kansalliskirjastossa on tapahtunut organisaatiouudistus ja PAS-palveluiden siirtyminen vapaakappaletoiimistoon, minkä takia asian eteneminen on hidastunut. Käytännön suunnittelu julkaisuarkistoista PASiin on käynnissä.

Aikataulu on seuraava: syksyllä aloitetaan käytännön työ ja putki olisi valmis ensi vuonna. Odotetaan PDF-muunnosasiassa ja mahdollisen Muuntajan käyttöönotossa syksyyn 2019, kun PASin tekniset vaatimukset varmistuvat.

Katkot

Katkoja on ollut jonkin verran. Pyydämme ilmoittamaan heti rukkaspostiin, jos katkoja havaitaan. Kansalliskirjastolla on hälytysjärjestelmä katkojen varalta ja he tiedottavat isommista katkoista Theseus.fi-etusivulla. Pääsääntöisesti tallennus on kuitenkin onnistunut, ellei opiskelija ole yrittänyt tehdä tallennusta Internet Explorer (IE) -selainta käyttämällä. Tallennussivulla on asiasta huomautus.

Ratkottuja ja työn alla olevia asioita

URN on saatu lähtemään opiskelijalle vastausviestissä. Palauteviestin kytkeminen suoraan tallennukseen ei ole tulossa lähiaikoina, mutta asiaa selvitetään.

Word-tiedostojen estoa tallennuksesta selvitetään parhaillaan. Ongelmana se, että Theseuksessa saattaa olla jo tallennettuna word-tiedostoja, joita ei voida eston jälkeen enää käyttää. Samoin selvitetään opiskelijan tallennuksen ohjaamista suoraan tiettyyn kokoelmaan Haka-tunnistuksen perusteella. Tämä vaatii Shibboleth-muutosta, tätä ratkotaan parhaillaan Aapa-verkoston kanssa.

Theseukseen mahdollisesti tarvittavia saavutettavuusdirektiivin mukaisia muutoksia selvitetään tänä vuonna niin ammattikorkeakoulujen käytänteinä kuin teknisen toteutuksen vaatimuksina. Ennen 23.9.2018 julkaistujen verkkosivustojen pitää olla saavutettavuusvaatimusten mukaisia 23.9.2020.

Striimaukseen liittyvät sopimusasiat ovat työn alla. Samoin julkaisujen pysyväistallennusasiaa otetaan esille Theseus-kokouksissa syksyllä, katsotaan, mitä asialle voi tehdä.

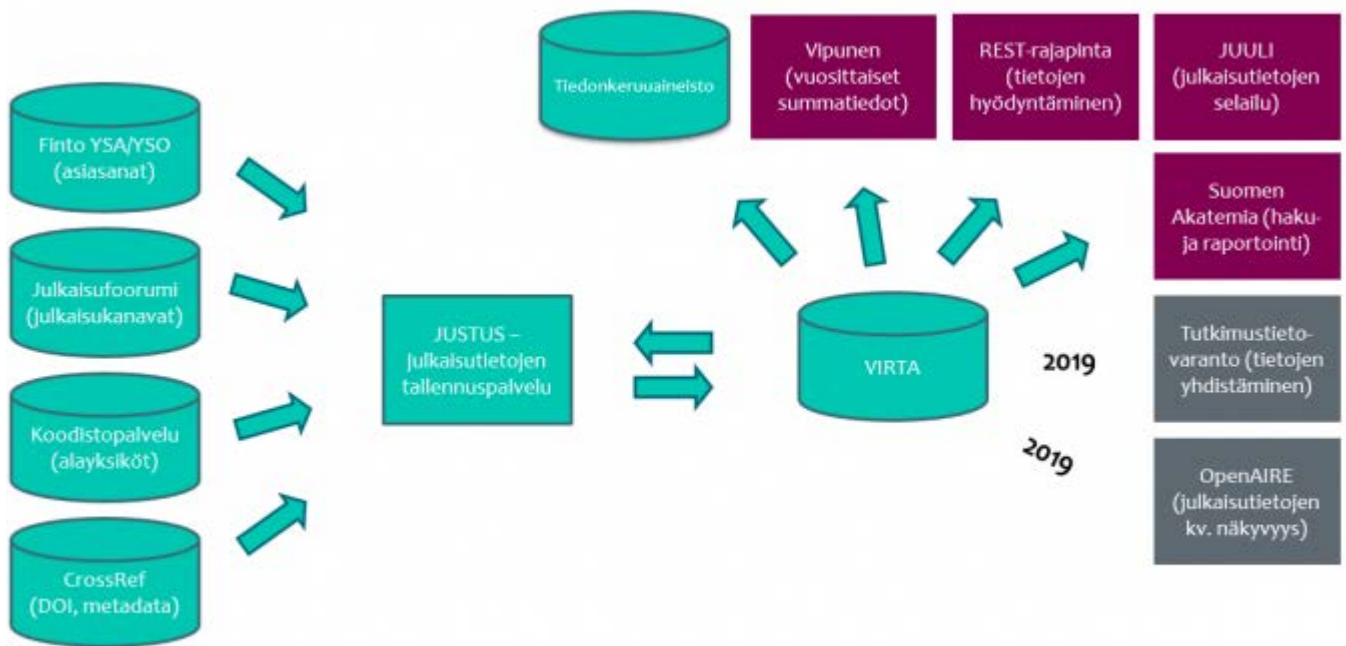
Viitetietoboxi

Tällä hetkellä Theseuksen viitetietoboxissa osa kentistä (esim. lehden nimeke -kenttä) eivät tule oikein näkyviin. Viitetietoboxia kehitetään yhteisessä kansallisessa työryhmässä, jossa tehtyjen päätösten jälkeen tehdään muutoksia Theseuksen boxin muodostavaan koodiin. Tämä menee syksylle.

Justusta hiotaan vielä

Justuksen kenttiin ja toiminnallisuuksiin liittyviä kommentteja otetaan vastaan. Välitämme kommentit CSC:lle ja he ovat aika nopeasti tarttuneet työhön.

Justuksessa sarjan/lehden nimeä ei pysty viemään eri kenttiin. Justus noudattaa ensisijaisesti VIRTAtietomallia ja [tiedonkeruun käsikirjan mukaisesti](#) lehden/sarjan nimi syötetään samaan kenttään. Tämän takia näiden erottaminen eri kenttiin ei toistaiseksi ole mahdollista Theseus-tiedonsiirrossa.



Justuksesta ammattikorkeakoulujen julkaisujen metadata siirtyy Virtaan ja sieltä eteenpäin metatietovarantoon, Vipuseen, Juuliin, Suomen Akatemialle ja OpenAIREen. © CSC

Tällä hetkellä selvitetään embargon, cc-lisenssin ja final draft/publishers edition -määrityksen poistamista peruskäyttäjältä kokonaan pääkäyttäjän tehtäväksi. Samoin rahoittaja -kentästä sekä tallentajan tiedoston takaisinlatausmahdollisuustoiveesta on raportoitu CSC:lle kuten toiveesta sarjan ja lehden nimekekkenttien eriyttämisestä. Rinnakkaistallennuslisenssin viemistä Theseuksen bundleen mietitään sekä korkeakoulun oman julkaisusarjaan liittyviä käytänteitä.

OpenAIRE-yhteys Theseuksen julkaisuille syntyy jatkossa Justuksen ja Virran kautta. Selvitämme syksyllä vielä lisää Theseukseen mahdollisesti otettavia OpenAIRE-metadatakenttiä, mutta voi olla, etteivät ne jatkossa ole tarpeen Virta-OpenAIRE-yhteyden takia.

Tietoa kirjoittajista:

Tiina Tolonen, palvelupäällikkö Oulun ammattikorkeakoulu
 Antti Nyqvist, tietopalvelusihteeri Haaga-Helia
 Minna Marjamaa, informaatikko Laurea